WEST Search History

DATE: Tuesday, October 15, 2002

Set Name side by side	Query	Hit Count	Set Name result set
	DB=USPT,PGPB,JPAB,EPAB,DWPI,TDBD; PLUR=YES; OP=OR		
L17	L16 and (solid phase synthes$4) adj4 (peptide or polypeptide or dimer or cycl$4)	18	L17
L16	L15 and (coupl$4 or bridg$4 or coupl$4) adj4 (agent or moiety)	445	L16
L15	L14 and iminodiacetic adj4 acid	2227	L15
L14	s (di or tri or tetra) adj4 carboxylic adj4 acid	5188532	L14
L13	L12 and iminodiacetic adj4 acid	54	L13
L12	L11 and (coup$6 or link$6) adj6 (agent or technique)	608	L12
L11	L10 and ring adj5 structure	2074	L11
L10	L5 and dimer$6	30286	L10
L9	L7 and achiral	1	L9
L8	L7 and imino adj4(di or tri or tetra)adj4 carboxylic adj4 acid	0	L8
L7	L6 and (aspartic or glutamic or glutaric)adj4 acid	525	L7
L6	L5 and iminodiacetic	1788	L6
L5	l1 and (solid phase synthesis)	868072	L5
L4	L2 and solid adj4 phase adj4 synthesis	26	L4
L3	L2 and achiral	8	L3
L2	L1 and (di or tri or tetra)adj4 carboxylic adj4 acid	9742	L2
L1	s borrelia	5179361	L1

END OF SEARCH HISTORY

```
     FILE 'CA' ENTERED AT 11:01:02 ON 15 OCT 2002
L1          594 S (DI OR TRI OR TETRA)(5W)CARBOXY?(5W)ACID
L2            1 S L1 AND BORRELIA BURGDORFERI
L3            1 S L1 AND ACHIRAL
L4            0 S L1 NOT L2 AND L3
L5          594 S L1 NOT L2 OR L3
L6          593 S L1 NOT L2
L7          593 S L6 NOT L3

     FILE 'REGISTRY' ENTERED AT 11:04:43 ON 15 OCT 2002
L8            7 S PVVAESPKKP/SQEP

     FILE 'CA' ENTERED AT 11:05:45 ON 15 OCT 2002
L9            5 S L8
L10           1 S L1 AND BORRELIA
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: October 12, 2002, 20:43:46 ; Search time 30 Seconds
(without alignments)
37.025 Million cell updates/sec

Title: US-09-408-578A-1
Perfect score: 52
Sequence: 1 PVVAESPKKP 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 352077

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802:*
1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	52	100.0	10	18	AAW41844	Modified B. burgdo
2	52	100.0	10	18	AAW41821	B. burgdorferi sen
3	52	100.0	11	18	AAW41825	Modified B. burgdo
4	52	100.0	15	18	AAW41827	B. burgdorferi sen
5	52	100.0	20	18	AAW41826	B. burgdorferi sen
6	52	100.0	23	20	AAY27428	Borrelia outer sur
7	52	100.0	24	20	AAY27429	Borrelia outer sur
8	49	94.2	10	18	AAW41838	Modified B. burgdo
9	49	94.2	11	15	AAR70367	Borrelia OspC anti
10	49	94.2	15	15	AAR70362	Borrelia OspC anti
11	48	92.3	10	18	AAW41841	Modified B. burgdo
12	48	92.3	10	18	AAW41842	Modified B. burgdo
13	47	90.4	10	18	AAW41848	Modified B. burgdo
14	47	90.4	10	18	AAW41849	Modified B. burgdo
15	46	88.5	10	18	AAW41845	Modified B. burgdo
16	46	88.5	10	18	AAW41835	Modified B. burgdo
17	46	88.5	10	18	AAW41836	Modified B. burgdo
18	46	88.5	10	18	AAW41839	Modified B. burgdo
19	46	88.5	10	18	AAW41840	Modified B. burgdo
20	45	86.5	9	18	AAW41828	B. burgdorferi sen
21	45	86.5	10	15	AAR70369	Borrelia OspC anti
22	45	86.5	10	18	AAW41843	Modified B. burgdo
23	45	86.5	10	18	AAW41834	Modified B. burgdo
24	44	84.6	10	18	AAW41837	Modified B. burgdo
25	43	82.7	10	18	AAW41846	Modified B. burgdo
26	41	78.8	8	18	AAW41829	B. burgdorferi sen
27	41	78.8	8	20	AAY27427	Borrelia outer sur
28	37	71.2	7	18	AAW41830	B. burgdorferi sen
29	35	67.3	10	18	AAW41850	Modified B. burgdo
30	34	65.4	10	22	AAM98380	Human peptide #165
31	33	63.5	6	18	AAW41831	B. burgdorferi sen
32	32	61.5	13	20	AAY41944	Rheumatoid arthrit
33	32	61.5	13	22	ABB52107	Human API-72 trypt
34	31	59.6	15	22	AAB86802	Human cytomegalovi
35	31	59.6	15	22	AAB86803	Human cytomegalovi
36	31	59.6	16	19	AAW80743	Mouse HDGF nucleus
37	31	59.6	24	21	AAB30413	Nuclear localisati
38	31	59.6	38	22	AAM00851	Human bone marrow
39	31	59.6	45	22	AAO04398	Human polypeptide
40	30.5	58.7	19	21	AAY85066	Immunogenic peptid
41	30	57.7	15	18	AAW29485	Peptide #6 derived
42	30	57.7	15	20	AAY34049	Histone H1S-3 deri
43	30	57.7	15	20	AAY34063	Histone H1 isoform
44	30	57.7	15	21	AAY57347	Human histone H1 p
45	30	57.7	15	21	AAY57361	Human histone H1.5

ALIGNMENTS

RESULT 1
AAW41844
ID AAW41844 standard; peptide; 10 AA.
XX
AC AAW41844;
XX
DT 14-MAY-1998 (first entry)
XX
DE Modified B. burgdorferi sensu lato OspC C-terminal peptide.
XX
KW Sensu lato; outer surface protein C; OspC; diagnosis; Lyme disease;
KW vaccine; infection.
XX
OS Borrelia burgdorferi.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 10
FT /label= amidated
XX
PN WO9742221-A1.
XX
PD 13-NOV-1997.
XX
PF 02-MAY-1997; 97WO-DK00203.
XX
PR 02-MAY-1996; 96DK-0000526.
XX
PA (STAT-) STATENS SERUMINSTITUT.
XX
PI Holm A, Mathiesen MJ, Ostergaard S, Theisen M;
XX
DR WPI; 1997-558908/51.
XX

PT Detecting previous sensitisation to the OspC protein of Borellia
PT burgdorferi - by detecting immunoreactivity between patient T cells
PT or immunoglobulins and C-terminal peptide of the protein
XX
PS Example 3; Page 53; 95pp; English.
XX
CC The present sequence was used in the development of a novel method
CC for the identification of a patient's previous sensitisation to
CC Borellia burgdorferi sensu lato outer surface protein C (OspC).
CC The method comprises reacting immunoglobulin (Ig) or T cells from
CC the patient with a polypeptide of at most 60 amino acids containing
CC a peptide with at least 50% identity to the B. burgdorferi derived
CC sequence AAW41821, or its subsequences of at least 5 amino acids. The
CC degree of immunological reactivity between the polypeptide and Ig
CC or T cells is measured and significant reactivity is indicative of
CC sensitisation.
CC The method can be used to diagnose Lyme disease and is based on
CC reactivity with antibodies against the OspC protein. The test can
CC be done in vitro or in vivo, e.g. as a skin test. Vaccine
CC compositions comprising the polypeptide can be used to protect
CC humans and other animals against B. burgdorferi infection. The
CC polypeptide provides higher sensitivity than full-length OspC, and
CC so is better at detecting infection in its early stages, especially
CC when combined with the known assay for flagellar proteins. The
CC seven carboxy-terminal residues of AAW41821 represent an epitope
CC essential for human immune response to OspC. The polypeptide is
CC also easier to prepare and purify than (nearly) full-length
CC protein, facilitating standardisation of the assay, and is less
CC cross-reactive with antibodies raised against other antigens. The
CC small size of the polypeptide allows a high density of binding
CC sites to be created on a solid support. Incorporation of
CC non-natural amino acid into the polypeptide increases its
CC resistance to peptidases when used in vivo.
XX
SQ Sequence 10 AA;

Query Match 100.0%; Score 52; DB 18; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0078;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PVVAESPKKP 10
||||||||||
Db 1 pvvaespkkp 10

RESULT 2
AAW41821
ID AAW41821 standard; peptide; 10 AA.
XX
AC AAW41821;
XX
DT 14-MAY-1998 (first entry)
XX
DE B. burgdorferi sensu lato OspC carboxy-terminal peptide.
XX
KW Sensu lato; outer surface protein C; OspC; diagnosis; Lyme disease;
KW vaccine; infection.
XX
OS Borrelia burgdorferi.
XX
PN WO9742221-A1.
XX
PD 13-NOV-1997.
XX
PF 02-MAY-1997; 97WO-DK00203.
XX
PR 02-MAY-1996; 96DK-0000526.
XX
PA (STAT-) STATENS SERUMINSTITUT.
XX
PI Holm A, Mathiesen MJ, Ostergaard S, Theisen M;
XX
DR WPI; 1997-558908/51.
XX
PT Detecting previous sensitisation to the OspC protein of Borellia
PT burgdorferi - by detecting immunoreactivity between patient T cells
PT or immunoglobulins and C-terminal peptide of the protein
XX
PS Claim 1; Page 77; 95pp; English.
XX
CC The present sequence was used in the development of a novel method
CC for the identification of a patient's previous sensitisation to
CC Borellia burgdorferi sensu lato outer surface protein C (OspC).
CC The method comprises reacting immunoglobulin (Ig) or T cells from
CC the patient with a polypeptide of at most 60 amino acids containing
CC a peptide with at least 50% identity to the B. burgdorferi derived
CC sequence AAW41821, or its subsequences of at least 5 amino acids. The
CC degree of immunological reactivity between the polypeptide and Ig
CC or T cells is measured and significant reactivity is indicative of
CC sensitisation.
CC The method can be used to diagnose Lyme disease and is based on
CC reactivity with antibodies against the OspC protein. The test can
CC be done in vitro or in vivo, e.g. as a skin test. Vaccine
CC compositions comprising the polypeptide can be used to protect
CC humans and other animals against B. burgdorferi infection. The
CC polypeptide provides higher sensitivity than full-length OspC, and
CC so is better at detecting infection in its early stages, especially
CC when combined with the known assay for flagellar proteins. The
CC seven carboxy-terminal residues of AAW41821 represent an epitope
CC essential for human immune response to OspC. The polypeptide is
CC also easier to prepare and purify than (nearly) full-length
CC protein, facilitating standardisation of the assay, and is less
CC cross-reactive with antibodies raised against other antigens. The
CC small size of the polypeptide allows a high density of binding
CC sites to be created on a solid support. Incorporation of
CC non-natural amino acid into the polypeptide increases its
CC resistance to peptidases when used in vivo.
XX
SQ Sequence 10 AA;

Query Match 100.0%; Score 52; DB 18; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0078;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PVVAESPKKP 10
||||||||||
Db 1 pvvaespkkp 10

RESULT 3
AAW41825
ID AAW41825 standard; peptide; 11 AA.
XX
AC AAW41825;
XX
DT 14-MAY-1998 (first entry)
XX
DE Modified B. burgdorferi sensu lato OspC C-terminal peptide.
XX
KW Sensu lato; outer surface protein C; OspC; diagnosis; Lyme disease;
KW vaccine; infection.
XX
OS Borrelia burgdorferi.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /note= "6-aminohexanoic acid"
XX
PN WO9742221-A1.
XX
PD 13-NOV-1997.
XX
PF 02-MAY-1997; 97WO-DK00203.

XX
PR 02-MAY-1996; 96DK-0000526.
XX
PA (STAT-) STATENS SERUMINSTITUT.
XX
PI Holm A, Mathiesen MJ, Ostergaard S, Theisen M;
XX
DR WPI; 1997-558908/51.
XX
PT Detecting previous sensitisation to the OspC protein of Borellia
PT burgdorferi - by detecting immunoreactivity between patient T cells
PT or immunoglobulins and C-terminal peptide of the protein
XX
PS Example; Page 40; 95pp; English.
XX
CC The present sequence was used in the development of a novel method
CC for the identification of a patient's previous sensitisation to
CC Borellia burgdorferi sensu lato outer surface protein C (OspC).
CC The method comprises reacting immunoglobulin (Ig) or T cells from
CC the patient with a polypeptide of at most 60 amino acids containing
CC a peptide with at least 50% identity to the B. burgdorferi derived
CC sequence AAW41821, or its subsequences of at least 5 amino acids. The
CC degree of immunological reactivity between the polypeptide and Ig
CC or T cells is measured and significant reactivity is indicative of
CC sensitisation.
CC The method can be used to diagnose Lyme disease and is based on
CC reactivity with antibodies against the OspC protein. The test can
CC be done in vitro or in vivo, e.g. as a skin test. Vaccine
CC compositions comprising the polypeptide can be used to protect
CC humans and other animals against B. burgdorferi infection. The
CC polypeptide provides higher sensitivity than full-length OspC, and
CC so is better at detecting infection in its early stages, especially
CC when combined with the known assay for flagellar proteins. The
CC seven carboxy-terminal residues of AAW41821 represent an epitope
CC essential for human immune response to OspC. The polypeptide is
CC also easier to prepare and purify than (nearly) full-length
CC protein, facilitating standardisation of the assay, and is less
CC cross-reactive with antibodies raised against other antigens. The
CC small size of the polypeptide allows a high density of binding
CC sites to be created on a solid support. Incorporation of
CC non-natural amino acid into the polypeptide increases its
CC resistance to peptidases when used in vivo.
XX
SQ Sequence 11 AA;

Query Match 100.0%; Score 52; DB 18; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.0085;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PVVAESPKKP 10
||||||||||
Db 2 pvvaespkkp 11

RESULT 4
AAW41827
ID AAW41827 standard; peptide; 15 AA.
XX
AC AAW41827;
XX
DT 14-MAY-1998 (first entry)
XX
DE B. burgdorferi sensu lato OspC carboxy-terminal peptide.
XX
KW Sensu lato; outer surface protein C; OspC; diagnosis; Lyme disease;
KW vaccine; infection.
XX
OS Borrelia burgdorferi.
XX
PN WO9742221-A1.
XX
PD 13-NOV-1997.
XX
PF 02-MAY-1997; 97WO-DK00203.
XX
PR 02-MAY-1996; 96DK-0000526.
XX
PA (STAT-) STATENS SERUMINSTITUT.
XX
PI Holm A, Mathiesen MJ, Ostergaard S, Theisen M;
XX
DR WPI; 1997-558908/51.
XX
PT Detecting previous sensitisation to the OspC protein of Borellia
PT burgdorferi - by detecting immunoreactivity between patient T cells
PT or immunoglobulins and C-terminal peptide of the protein
XX
PS Example 3; Page 51; 95pp; English.
XX
CC The present sequence was used in the development of a novel method
CC for the identification of a patient's previous sensitisation to
CC Borellia burgdorferi sensu lato outer surface protein C (OspC).
CC The method comprises reacting immunoglobulin (Ig) or T cells from
CC the patient with a polypeptide of at most 60 amino acids containing
CC a peptide with at least 50% identity to the B. burgdorferi derived
CC sequence AAW41821, or its subsequences of at least 5 amino acids. The
CC degree of immunological reactivity between the polypeptide and Ig
CC or T cells is measured and significant reactivity is indicative of
CC sensitisation.
CC The method can be used to diagnose Lyme disease and is based on
CC reactivity with antibodies against the OspC protein. The test can
CC be done in vitro or in vivo, e.g. as a skin test. Vaccine
CC compositions comprising the polypeptide can be used to protect
CC humans and other animals against B. burgdorferi infection. The
CC polypeptide provides higher sensitivity than full-length OspC, and
CC so is better at detecting infection in its early stages, especially
CC when combined with the known assay for flagellar proteins. The
CC seven carboxy-terminal residues of AAW41821 represent an epitope
CC essential for human immune response to OspC. The polypeptide is
CC also easier to prepare and purify than (nearly) full-length
CC protein, facilitating standardisation of the assay, and is less
CC cross-reactive with antibodies raised against other antigens. The
CC small size of the polypeptide allows a high density of binding
CC sites to be created on a solid support. Incorporation of
CC non-natural amino acid into the polypeptide increases its
CC resistance to peptidases when used in vivo.
XX
SQ Sequence 15 AA;

Query Match 100.0%; Score 52; DB 18; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.012;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PVVAESPKKP 10
||||||||||
Db 6 pvvaespkkp 15

RESULT 5
AAW41826
ID AAW41826 standard; peptide; 20 AA.
XX
AC AAW41826;
XX
DT 14-MAY-1998 (first entry)
XX
DE B. burgdorferi sensu lato OspC carboxy-terminal peptide.
XX
KW Sensu lato; outer surface protein C; OspC; diagnosis; Lyme disease;
KW vaccine; infection.
XX
OS Borrelia burgdorferi.
XX
PN WO9742221-A1.

```
XX
PD   13-NOV-1997.
XX
PF   02-MAY-1997;   97WO-DK00203.
XX
PR   02-MAY-1996;   96DK-0000526.
XX
PA   (STAT-) STATENS SERUMINSTITUT.
XX
PI   Holm A,  Mathiesen MJ,  Ostergaard S,  Theisen M;
XX
DR   WPI; 1997-558908/51.
XX
PT   Detecting previous sensitisation to the OspC protein of Borellia
PT   burgdorferi - by detecting immunoreactivity between patient T cells
PT   or immunoglobulins and C-terminal peptide of the protein
XX
PS   Example 3; Page 51; 95pp; English.
XX
CC   The present sequence was used in the development of a novel method
CC   for the identification of a patient's previous sensitisation to
CC   Borellia burgdorferi sensu lato outer surface protein C (OspC).
CC   The method comprises reacting immunoglobulin (Ig) or T cells from
CC   the patient with a polypeptide of at most 60 amino acids containing
CC   a peptide with at least 50% identity to the B. burgdorferi derived
CC   sequence AAW41821, or its subsequences of at least 5 amino acids. The
CC   degree of immunological reactivity between the polypeptide and Ig
CC   or T cells is measured and significant reactivity is indicative of
CC   sensitisation.
CC   The method can be used to diagnose Lyme disease and is based on
CC   reactivity with antibodies against the OspC protein. The test can
CC   be done in vitro or in vivo, e.g. as a skin test. Vaccine
CC   compositions comprising the polypeptide can be used to protect
CC   humans and other animals against B. burgdorferi infection. The
CC   polypeptide provides higher sensitivity than full-length OspC, and
CC   so is better at detecting infection in its early stages, especially
CC   when combined with the known assay for flagellar proteins. The
CC   seven carboxy-terminal residues of AAW41821 represent an epitope
CC   essential for human immune response to OspC. The polypeptide is
CC   also easier to prepare and purify than (nearly) full-length
CC   protein, facilitating standardisation of the assay, and is less
CC   cross-reactive with antibodies raised against other antigens. The
CC   small size of the polypeptide allows a high density of binding
CC   sites to be created on a solid support. Incorporation of
CC   non-natural amino acid into the polypeptide increases its
CC   resistance to peptidases when used in vivo.
XX
SQ   Sequence   20 AA;

  Query Match             100.0%;  Score 52;  DB 18;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 0.016;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 PVVAESPKKP 10
           ||||||||||
Db      11 pvvaespkkp 20


RESULT   6
AAY27428
ID   AAY27428 standard; peptide; 23 AA.
XX
AC   AAY27428;
XX
DT   26-NOV-1999  (first entry)
XX
DE   Borrelia outer surface protein C (OspC) C-terminal peptide fragment.
XX
KW   Borrelia; IgM antibody; outer surface protein C; OspC; deer tick;
KW   cerebrospinal fluid; Lyme borreliosis; micro-capture assay; flagellum;
KW   epitope.
XX
OS   Synthetic.
OS   Borrelia burgdorferi.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "linked to biotin via an O-linker of formula
FT                          [2-aminoethoxy]ethoxy acetic acid"
XX
PN   EP949508-A1.
XX
PD   13-OCT-1999.
XX
PF   07-APR-1999;   99EP-0610026.
XX
PR   08-APR-1998;   98DK-0000516.
XX
PA   (DAKO-) DAKO AS.
XX
PI   Staffeldt Schou O,  Winther L,  Stender H;
XX
DR   WPI; 1999-553537/47.
XX
PT   Diagnosing Lyme borreliosis by detecting antibodies against two
PT   antigens simultaneously -
XX
PS   Example 1; Page 7; 23pp; English.
XX
CC   The invention provides a new method for detecting IgM antibodies against
CC   Borrelia burgdorferi in a sample of human or animal fluid. The method
CC   comprises: (1) contacting antibodies in the sample with anti-IgM
CC   immobilized to a solid support, (2) separating the support from the
CC   liquid phase; and (3) contacting the bound antibodies with a complex
CC   comprising at least one set of B. burgdorferi (outer surface protein C
CC   (OspC) peptides and/or at least one set of other B. burgdorferi peptides,
CC   each attached to a carrier; and (4) detecting the presence of antibodies
CC   against B. burgdorferi. The new method may be used to detect antibodies
CC   against B. burgdorferi in (especially) serum or cerebrospinal fluid
CC   samples from patients bitten by deer ticks. B. burgdorferi causes Lyme
CC   borreliosis so detection of antibodies against it allows diagnosis of
CC   infection by this organism. The method is a micro-capture assay in which
CC   the antigen complex is a combination of the B.burgdorferi flagellum and
CC   OspC peptides. The presence of epitopes from both antigens in the complex
CC   allows the simultaneous detection of serum antibodies against these
CC   proteins which increases the sensitivity of the test. The two antigens
CC   are pure, which also decreases the possibility of cross-reactivity. The
CC   present sequence represents a Borrelia OspC C-terminal peptide fragment.
XX
SQ   Sequence   23 AA;

  Query Match             100.0%;  Score 52;  DB 20;  Length 23;
  Best Local Similarity   100.0%;  Pred. No. 0.018;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 PVVAESPKKP 10
           ||||||||||
Db      14 pvvaespkkp 23


RESULT   7
AAY27429
ID   AAY27429 standard; peptide; 24 AA.
XX
AC   AAY27429;
XX
DT   26-NOV-1999  (first entry)
XX
DE   Borrelia outer surface protein C (OspC) C-terminal peptide fragment.
XX
KW   Borrelia; IgM antibody; outer surface protein C; OspC; deer tick;
KW   cerebrospinal fluid; Lyme borreliosis; micro-capture assay; flagellum;
KW   epitope.
XX
```

OS Synthetic.
OS Borrelia burgdorferi.
XX
PN EP949508-A1.
XX
PD 13-OCT-1999.
XX
PF 07-APR-1999; 99EP-0610026.
XX
PR 08-APR-1998; 98DK-0000516.
XX
PA (DAKO-) DAKO AS.
XX
PI Staffeldt Schou O, Winther L, Stender H;
XX
DR WPI; 1999-553537/47.
XX
PT Diagnosing Lyme borreliosis by detecting antibodies against two
PT antigens simultaneously -
XX
PS Example 1; Page 7; 23pp; English.
XX
CC The invention provides a new method for detecting IgM antibodies against
CC Borrelia burgdorferi in a sample of human or animal fluid. The method
CC comprises: (1) contacting antibodies in the sample with anti-IgM
CC immobilized to a solid support, (2) separating the support from the
CC liquid phase; and (3) contacting the bound antibodies with a complex
CC comprising at least one set of B. burgdorferi (outer surface protein C
CC (OspC) peptides and/or at least one set of other B. burgdorferi peptides,
CC each attached to a carrier; and (4) detecting the presence of antibodies
CC against B. burgdorferi. The new method may be used to detect antibodies
CC against B. burgdorferi in (especially) serum or cerebrospinal fluid
CC samples from patients bitten by deer ticks. B. burgdorferi causes Lyme
CC borreliosis so detection of antibodies against it allows diagnosis of
CC infection by this organism. The method is a micro-capture assay in which
CC the antigen complex is a combination of the B.burgdorferi flagellum and
CC OspC peptides. The presence of epitopes from both antigens in the complex
CC allows the simultaneous detection of serum antibodies against these
CC proteins which increases the sensitivity of the test. The two antigens
CC are pure, which also decreases the possibility of cross-reactivity. The
CC present sequence represents a Borrelia OspC C-terminal peptide fragment,
CC where the N-terminal cysteine residue has been incorporated to provide a
CC SH group to be used in a coupling reaction.
XX
SQ Sequence 24 AA;

Query Match 100.0%; Score 52; DB 20; Length 24;
Best Local Similarity 100.0%; Pred. No. 0.019;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PVVAESPKKP 10
||||||||||
Db 15 pvvaespkkp 24

RESULT 8
AAW41838
ID AAW41838 standard; peptide; 10 AA.
XX
AC AAW41838;
XX
DT 14-MAY-1998 (first entry)
XX
DE Modified B. burgdorferi sensu lato OspC C-terminal peptide.
XX
KW Sensu lato; outer surface protein C; OspC; diagnosis; Lyme disease;
KW vaccine; infection.
XX
OS Borrelia burgdorferi.
OS Synthetic.
XX
PN WO9742221-A1.
XX
PD 13-NOV-1997.
XX
PF 02-MAY-1997; 97WO-DK00203.
XX
PR 02-MAY-1996; 96DK-0000526.
XX
PA (STAT-) STATENS SERUMINSTITUT.
XX
PI Holm A, Mathiesen MJ, Ostergaard S, Theisen M;
XX
DR WPI; 1997-558908/51.
XX
PT Detecting previous sensitisation to the OspC protein of Borellia
PT burgdorferi - by detecting immunoreactivity between patient T cells
PT or immunoglobulins and C-terminal peptide of the protein
XX
PS Example 3; Page 52; 95pp; English.
XX
CC The present sequence was used in the development of a novel method
CC for the identification of a patient's previous sensitisation to
CC Borellia burgdorferi sensu lato outer surface protein C (OspC).
CC The method comprises reacting immunoglobulin (Ig) or T cells from
CC the patient with a polypeptide of at most 60 amino acids containing
CC a peptide with at least 50% identity to the B. burgdorferi derived
CC sequence AAW41821, or its subsequences of at least 5 amino acids. The
CC degree of immunological reactivity between the polypeptide and Ig
CC or T cells is measured and significant reactivity is indicative of
CC sensitisation.
CC The method can be used to diagnose Lyme disease and is based on
CC reactivity with antibodies against the OspC protein. The test can
CC be done in vitro or in vivo, e.g. as a skin test. Vaccine
CC compositions comprising the polypeptide can be used to protect
CC humans and other animals against B. burgdorferi infection. The
CC polypeptide provides higher sensitivity than full-length OspC, and
CC so is better at detecting infection in its early stages, especially
CC when combined with the known assay for flagellar proteins. The
CC seven carboxy-terminal residues of AAW41821 represent an epitope
CC essential for human immune response to OspC. The polypeptide is
CC also easier to prepare and purify than (nearly) full-length
CC protein, facilitating standardisation of the assay, and is less
CC cross-reactive with antibodies raised against other antigens. The
CC small size of the polypeptide allows a high density of binding
CC sites to be created on a solid support. Incorporation of
CC non-natural amino acid into the polypeptide increases its
CC resistance to peptidases when used in vivo.
XX
SQ Sequence 10 AA;

Query Match 94.2%; Score 49; DB 18; Length 10;
Best Local Similarity 90.0%; Pred. No. 0.026;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PVVAESPKKP 10
||||| ||||
Db 1 pvvaeapkkp 10

RESULT 9
AAR70367
ID AAR70367 standard; peptide; 11 AA.
XX
AC AAR70367;
XX
DT 25-MAY-1995 (first entry)
XX
DE Borrelia OspC antigen epitope.
XX
KW OspC antigen; vaccine; Lyme disease; borreliosis; immunogen;
KW serovar typing; restriction fragment length polymorphism;
KW RFLP.
XX

OS Borrelia burgdorferi.
XX
PN WO9425596-A.
XX
PD 10-NOV-1994.
XX
PF 29-APR-1994; 94WO-EP01365.
XX
PR 29-APR-1993; 93US-0053863.
XX
PA (IMMO) IMMUNO AG.
XX
PI Crowe B, Dorner F, Livey I;
XX
DR WPI; 1994-358273/44.
XX
PT Immunogenic composition comprising OspC antigens - for the
PT treatment of Lyme borreliosis in different, specific geographical
PT areas.
XX
PS Claim 19; Page 57; 115pp; English.
XX
CC A vaccine for Lyme disease includes selected OspC antigen
CC formulations based on defined OspC families resolved by serovar
CC typing and RFLP typing of strains of worldwide origin. The
CC antigens comprise 1 or more of the epitopes given in AAR70360-69
CC or their variants or mimetics.
XX
SQ Sequence 11 AA;

Query Match 94.2%; Score 49; DB 15; Length 11;
Best Local Similarity 90.0%; Pred. No. 0.029;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PVVAESPKKP 10
|||||:||||
Db 2 pvvaetpkkp 11

RESULT 10
AAR70362
ID AAR70362 standard; peptide; 15 AA.
XX
AC AAR70362;
XX
DT 25-MAY-1995 (first entry)
XX
DE Borrelia OspC antigen epitope.
XX
KW OspC antigen; vaccine; Lyme disease; borreliosis; immunogen;
KW serovar typing; restriction fragment length polymorphism;
KW RFLP.
XX
OS Borrelia burgdorferi.
XX
PN WO9425596-A.
XX
PD 10-NOV-1994.
XX
PF 29-APR-1994; 94WO-EP01365.
XX
PR 29-APR-1993; 93US-0053863.
XX
PA (IMMO) IMMUNO AG.
XX
PI Crowe B, Dorner F, Livey I;
XX
DR WPI; 1994-358273/44.
XX
PT Immunogenic composition comprising OspC antigens - for the
PT treatment of Lyme borreliosis in different, specific geographical
PT areas.
XX
PS Claim 19; Page 56; 115pp; English.
XX
CC A vaccine for Lyme disease includes selected OspC antigen
CC formulations based on defined OspC families resolved by serovar
CC typing and RFLP typing of strains of worldwide origin. The
CC antigens comprise 1 or more of the epitopes given in AAR70360-69
CC or their variants or mimetics.
XX
SQ Sequence 15 AA;

Query Match 94.2%; Score 49; DB 15; Length 15;
Best Local Similarity 90.0%; Pred. No. 0.039;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PVVAESPKKP 10
|||||:||||
Db 6 pvvaetpkkp 15

RESULT 11
AAW41841
ID AAW41841 standard; peptide; 10 AA.
XX
AC AAW41841;
XX
DT 14-MAY-1998 (first entry)
XX
DE Modified B. burgdorferi sensu lato OspC C-terminal peptide.
XX
KW Sensu lato; outer surface protein C; OspC; diagnosis; Lyme disease;
KW vaccine; infection.
XX
OS Borrelia burgdorferi.
OS Synthetic.
XX
PN WO9742221-A1.
XX
PD 13-NOV-1997.
XX
PF 02-MAY-1997; 97WO-DK00203.
XX
PR 02-MAY-1996; 96DK-0000526.
XX
PA (STAT-) STATENS SERUMINSTITUT.
XX
PI Holm A, Mathiesen MJ, Ostergaard S, Theisen M;
XX
DR WPI; 1997-558908/51.
XX
PT Detecting previous sensitisation to the OspC protein of Borellia
PT burgdorferi - by detecting immunoreactivity between patient T cells
PT or immunoglobulins and C-terminal peptide of the protein
XX
PS Example 3; Page 52; 95pp; English.
XX
CC The present sequence was used in the development of a novel method
CC for the identification of a patient's previous sensitisation to
CC Borellia burgdorferi sensu lato outer surface protein C (OspC).
CC The method comprises reacting immunoglobulin (Ig) or T cells from
CC the patient with a polypeptide of at most 60 amino acids containing
CC a peptide with at least 50% identity to the B. burgdorferi derived
CC sequence AAW41821, or its subsequences of at least 5 amino acids. The
CC degree of immunological reactivity between the polypeptide and Ig
CC or T cells is measured and significant reactivity is indicative of
CC sensitisation.
CC The method can be used to diagnose Lyme disease and is based on
CC reactivity with antibodies against the OspC protein. The test can
CC be done in vitro or in vivo, e.g. as a skin test. Vaccine
CC compositions comprising the polypeptide can be used to protect
CC humans and other animals against B. burgdorferi infection. The
CC polypeptide provides higher sensitivity than full-length OspC, and

CC so is better at detecting infection in its early stages, especially
CC when combined with the known assay for flagellar proteins. The
CC seven carboxy-terminal residues of AAW41821 represent an epitope
CC essential for human immune response to OspC. The polypeptide is
CC also easier to prepare and purify than (nearly) full-length
CC protein, facilitating standardisation of the assay, and is less
CC cross-reactive with antibodies raised against other antigens. The
CC small size of the polypeptide allows a high density of binding
CC sites to be created on a solid support. Incorporation of
CC non-natural amino acid into the polypeptide increases its
CC resistance to peptidases when used in vivo.
XX
SQ Sequence 10 AA;

Query Match 92.3%; Score 48; DB 18; Length 10;
Best Local Similarity 90.0%; Pred. No. 0.039;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PVVAESPKKP 10
|| |||||||
Db 1 pvaaespkkp 10

RESULT 12
AAW41842
ID AAW41842 standard; peptide; 10 AA.
XX
AC AAW41842;
XX
DT 14-MAY-1998 (first entry)
XX
DE Modified B. burgdorferi sensu lato OspC C-terminal peptide.
XX
KW Sensu lato; outer surface protein C; OspC; diagnosis; Lyme disease;
KW vaccine; infection.
XX
OS Borrelia burgdorferi.
OS Synthetic.
XX
PN WO9742221-A1.
XX
PD 13-NOV-1997.
XX
PF 02-MAY-1997; 97WO-DK00203.
XX
PR 02-MAY-1996; 96DK-0000526.
XX
PA (STAT-) STATENS SERUMINSTITUT.
XX
PI Holm A, Mathiesen MJ, Ostergaard S, Theisen M;
XX
DR WPI; 1997-558908/51.
XX
PT Detecting previous sensitisation to the OspC protein of Borellia
PT burgdorferi - by detecting immunoreactivity between patient T cells
PT or immunoglobulins and C-terminal peptide of the protein
XX
PS Example 3; Page 52; 95pp; English.
XX
CC The present sequence was used in the development of a novel method
CC for the identification of a patient's previous sensitisation to
CC Borellia burgdorferi sensu lato outer surface protein C (OspC).
CC The method comprises reacting immunoglobulin (Ig) or T cells from
CC the patient with a polypeptide of at most 60 amino acids containing
CC a peptide with at least 50% identity to the B. burgdorferi derived
CC sequence AAW41821, or its subsequences of at least 5 amino acids. The
CC degree of immunological reactivity between the polypeptide and Ig
CC or T cells is measured and significant reactivity is indicative of
CC sensitisation.
CC The method can be used to diagnose Lyme disease and is based on
CC reactivity with antibodies against the OspC protein. The test can
CC be done in vitro or in vivo, e.g. as a skin test. Vaccine
CC compositions comprising the polypeptide can be used to protect
CC humans and other animals against B. burgdorferi infection. The
CC polypeptide provides higher sensitivity than full-length OspC, and
CC so is better at detecting infection in its early stages, especially
CC when combined with the known assay for flagellar proteins. The
CC seven carboxy-terminal residues of AAW41821 represent an epitope
CC essential for human immune response to OspC. The polypeptide is
CC also easier to prepare and purify than (nearly) full-length
CC protein, facilitating standardisation of the assay, and is less
CC cross-reactive with antibodies raised against other antigens. The
CC small size of the polypeptide allows a high density of binding
CC sites to be created on a solid support. Incorporation of
CC non-natural amino acid into the polypeptide increases its
CC resistance to peptidases when used in vivo.
XX
SQ Sequence 10 AA;

Query Match 92.3%; Score 48; DB 18; Length 10;
Best Local Similarity 90.0%; Pred. No. 0.039;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PVVAESPKKP 10
| ||||||||
Db 1 pavaespkkp 10

RESULT 13
AAW41848
ID AAW41848 standard; peptide; 10 AA.
XX
AC AAW41848;
XX
DT 14-MAY-1998 (first entry)
XX
DE Modified B. burgdorferi sensu lato OspC C-terminal peptide.
XX
KW Sensu lato; outer surface protein C; OspC; diagnosis; Lyme disease;
KW vaccine; infection.
XX
OS Borrelia burgdorferi.
OS Synthetic.
XX
PN WO9742221-A1.
XX
PD 13-NOV-1997.
XX
PF 02-MAY-1997; 97WO-DK00203.
XX
PR 02-MAY-1996; 96DK-0000526.
XX
PA (STAT-) STATENS SERUMINSTITUT.
XX
PI Holm A, Mathiesen MJ, Ostergaard S, Theisen M;
XX
DR WPI; 1997-558908/51.
XX
PT Detecting previous sensitisation to the OspC protein of Borellia
PT burgdorferi - by detecting immunoreactivity between patient T cells
PT or immunoglobulins and C-terminal peptide of the protein
XX
PS Example 3; Page 54; 95pp; English.
XX
CC The present sequence was used in the development of a novel method
CC for the identification of a patient's previous sensitisation to
CC Borellia burgdorferi sensu lato outer surface protein C (OspC).
CC The method comprises reacting immunoglobulin (Ig) or T cells from
CC the patient with a polypeptide of at most 60 amino acids containing
CC a peptide with at least 50% identity to the B. burgdorferi derived
CC sequence AAW41821, or its subsequences of at least 5 amino acids. The
CC degree of immunological reactivity between the polypeptide and Ig
CC or T cells is measured and significant reactivity is indicative of
CC sensitisation.

CC The method can be used to diagnose Lyme disease and is based on
CC reactivity with antibodies against the OspC protein. The test can
CC be done in vitro or in vivo, e.g. as a skin test. Vaccine
CC compositions comprising the polypeptide can be used to protect
CC humans and other animals against B. burgdorferi infection. The
CC polypeptide provides higher sensitivity than full-length OspC, and
CC so is better at detecting infection in its early stages, especially
CC when combined with the known assay for flagellar proteins. The
CC seven carboxy-terminal residues of AAW41821 represent an epitope
CC essential for human immune response to OspC. The polypeptide is
CC also easier to prepare and purify than (nearly) full-length
CC protein, facilitating standardisation of the assay, and is less
CC cross-reactive with antibodies raised against other antigens. The
CC small size of the polypeptide allows a high density of binding
CC sites to be created on a solid support. Incorporation of
CC non-natural amino acid into the polypeptide increases its
CC resistance to peptidases when used in vivo.
XX
SQ Sequence 10 AA;

Query Match 90.4%; Score 47; DB 18; Length 10;
Best Local Similarity 90.0%; Pred. No. 0.059;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PVVAESPKKP 10
||| ||||||
Db 1 pvvpespkkp 10

RESULT 14
AAW41849
ID AAW41849 standard; peptide; 10 AA.
XX
AC AAW41849;
XX
DT 14-MAY-1998 (first entry)
XX
DE Modified B. burgdorferi sensu lato OspC C-terminal peptide.
XX
KW Sensu lato; outer surface protein C; OspC; diagnosis; Lyme disease;
KW vaccine; infection.
XX
OS Borrelia burgdorferi.
OS Synthetic.
XX
PN WO9742221-A1.
XX
PD 13-NOV-1997.
XX
PF 02-MAY-1997; 97WO-DK00203.
XX
PR 02-MAY-1996; 96DK-0000526.
XX
PA (STAT-) STATENS SERUMINSTITUT.
XX
PI Holm A, Mathiesen MJ, Ostergaard S, Theisen M;
XX
DR WPI; 1997-558908/51.
XX
PT Detecting previous sensitisation to the OspC protein of Borellia
PT burgdorferi - by detecting immunoreactivity between patient T cells
PT or immunoglobulins and C-terminal peptide of the protein
XX
PS Example 3; Page 54; 95pp; English.
XX
CC The present sequence was used in the development of a novel method
CC for the identification of a patient's previous sensitisation to
CC Borellia burgdorferi sensu lato outer surface protein C (OspC).
CC The method comprises reacting immunoglobulin (Ig) or T cells from
CC the patient with a polypeptide of at most 60 amino acids containing
CC a peptide with at least 50% identity to the B. burgdorferi derived
CC sequence AAW41821, or its subsequences of at least 5 amino acids. The
CC degree of immunological reactivity between the polypeptide and Ig
CC or T cells is measured and significant reactivity is indicative of
CC sensitisation.
CC The method can be used to diagnose Lyme disease and is based on
CC reactivity with antibodies against the OspC protein. The test can
CC be done in vitro or in vivo; e.g. as a skin test. Vaccine
CC compositions comprising the polypeptide can be used to protect
CC humans and other animals against B. burgdorferi infection. The
CC polypeptide provides higher sensitivity than full-length OspC, and
CC so is better at detecting infection in its early stages, especially
CC when combined with the known assay for flagellar proteins. The
CC seven carboxy-terminal residues of AAW41821 represent an epitope
CC essential for human immune response to OspC. The polypeptide is
CC also easier to prepare and purify than (nearly) full-length
CC protein, facilitating standardisation of the assay, and is less
CC cross-reactive with antibodies raised against other antigens. The
CC small size of the polypeptide allows a high density of binding
CC sites to be created on a solid support. Incorporation of
CC non-natural amino acid into the polypeptide increases its
CC resistance to peptidases when used in vivo.
XX
SQ Sequence 10 AA;

Query Match 90.4%; Score 47; DB 18; Length 10;
Best Local Similarity 90.0%; Pred. No. 0.059;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PVVAESPKKP 10
||||||| |
Db 1 pvvaespknp 10

RESULT 15
AAW41845
ID AAW41845 standard; peptide; 10 AA.
XX
AC AAW41845;
XX
DT 14-MAY-1998 (first entry)
XX
DE Modified B. burgdorferi sensu lato OspC C-terminal peptide.
XX
KW Sensu lato; outer surface protein C; OspC; diagnosis; Lyme disease;
KW vaccine; infection.
XX
OS Borrelia burgdorferi.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 9
FT /label= Dpr, Dbu, Orn, Arg, Lys, OTHER
FT /note= "OTHER = diaminoacetic acid"
FT Arg and Lys are D-form residues"
XX
PN WO9742221-A1.
XX
PD 13-NOV-1997.
XX
PF 02-MAY-1997; 97WO-DK00203.
XX
PR 02-MAY-1996; 96DK-0000526.
XX
PA (STAT-) STATENS SERUMINSTITUT.
XX
PI Holm A, Mathiesen MJ, Ostergaard S, Theisen M;
XX
DR WPI; 1997-558908/51.
XX
PT Detecting previous sensitisation to the OspC protein of Borellia
PT burgdorferi - by detecting immunoreactivity between patient T cells
PT or immunoglobulins and C-terminal peptide of the protein
XX

PS Example 3; Page 53; 95pp; English.
XX
CC The present sequence was used in the development of a novel method
CC for the identification of a patient's previous sensitisation to
CC Borellia burgdorferi sensu lato outer surface protein C (OspC).
CC The method comprises reacting immunoglobulin (Ig) or T cells from
CC the patient with a polypeptide of at most 60 amino acids containing
CC a peptide with at least 50% identity to the B. burgdorferi derived
CC sequence AAW41821, or its subsequences of at least 5 amino acids. The
CC degree of immunological reactivity between the polypeptide and Ig
CC or T cells is measured and significant reactivity is indicative of
CC sensitisation.
CC The method can be used to diagnose Lyme disease and is based on
CC reactivity with antibodies against the OspC protein. The test can
CC be done in vitro or in vivo, e.g. as a skin test. Vaccine
CC compositions comprising the polypeptide can be used to protect
CC humans and other animals against B. burgdorferi infection. The
CC polypeptide provides higher sensitivity than full-length OspC, and
CC so is better at detecting infection in its early stages, especially
CC when combined with the known assay for flagellar proteins. The
CC seven carboxy-terminal residues of AAW41821 represent an epitope
CC essential for human immune response to OspC. The polypeptide is
CC also easier to prepare and purify than (nearly) full-length
CC protein, facilitating standardisation of the assay, and is less
CC cross-reactive with antibodies raised against other antigens. The
CC small size of the polypeptide allows a high density of binding
CC sites to be created on a solid support. Incorporation of
CC non-natural amino acid into the polypeptide increases its
CC resistance to peptidases when used in vivo.
XX
SQ Sequence 10 AA;

Query Match 88.5%; Score 46; DB 18; Length 10;
Best Local Similarity 90.0%; Pred. No. 0.089;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PVVAESPKKP 10
||||||||| |
Db 1 pvvaespkxp 10

Search completed: October 12, 2002, 20:48:43
Job time: 297 sec

GenCore version 4.5

OM protein - protein search, using sw model

Run on: October 12, 2002, 20:45:36 ; Search time 19.91 Seconds
(without alignments)
48.262 Million cell updates/sec

Title: US-09-408-578A-1
Perfect score: 52
Sequence: 1 PVVAESPKKP 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 11821

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	27	51.9	15	2	PS0251	15K protein 5106 -
2	26	50.0	14	2	PN0666	dystrophin-associa
3	26	50.0	32	2	S01810	hemoglobin AIV - t
4	26	50.0	47	2	E47395	histone H1 II-1 (c
5	26	50.0	47	2	S14064	hypothetical prote
6	25	48.1	20	2	I53671	neurofilament heav
7	25	48.1	30	2	S59482	hydroxyproline-ric
8	25	48.1	36	2	A82208	hypothetical prote
9	25	48.1	37	1	R5BS36	ribosomal protein
10	25	48.1	37	2	B32307	ribosomal protein
11	25	48.1	37	2	T44407	ribosomal protein
12	25	48.1	37	2	E97282	ribosomal protein
13	25	48.1	37	2	F90019	50S ribosomal prot
14	25	48.1	37	2	AH1776	ribosomal protein
15	25	48.1	37	2	AI1400	ribosomal protein
16	25	48.1	39	2	S16978	4.8K protein - Ana
17	25	48.1	40	2	A28938	polymorphic epithe
18	25	48.1	49	2	D88533	5.0K hypothetical
19	25	48.1	50	2	C88533	hypothetical prote
20	24	46.2	12	2	B39690	neural cell adhesi
21	24	46.2	17	2	S66198	alcohol dehydrogen
22	24	46.2	27	2	A30323	amyloid protein AL
23	24	46.2	37	2	E81254	50S ribosomal prot
24	24	46.2	37	2	G95010	hypothetical prote
25	24	46.2	39	2	T42897	hypothetical prote
26	24	46.2	41	2	S10263	histone H3.2 - Tet
27	24	46.2	42	2	T36992	probable transposa
28	24	46.2	45	2	S04941	protamine phi-3.1
29	24	46.2	45	2	S10544	protamine phi-3.2
30	24	46.2	45	2	S10545	protamine phi-3.3
31	24	46.2	48	2	I61693	myosin - human (fr
32	24	46.2	49	2	PC2062	dihemic cytochrome
33	24	46.2	49	2	A81605	hypothetical prote
34	24	46.2	50	2	D90706	hypothetical prote
35	24	46.2	50	2	F82409	hypothetical prote
36	23	44.2	11	2	PQ0231	beta-glucosidase (
37	23	44.2	15	2	A54397	ubiquitin-carrier
38	23	44.2	26	2	S06675	apidaecin Ib precu
39	23	44.2	27	2	S71302	ICL6 protein - Par
40	23	44.2	30	2	S72626	small-cell-variant
41	23	44.2	31	2	S50903	fatty acid beta-ox
42	23	44.2	31	2	T36103	hypothetical prote
43	23	44.2	32	2	C61491	seed protein ws-3
44	23	44.2	34	2	S17644	alcohol dehydrogen
45	23	44.2	36	2	A69287	hypothetical prote

ALIGNMENTS

RESULT 1
PS0251
15K protein 5106 - rice (strain Nihonbare) (fragment)
C;Species: Oryza sativa (rice)
C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 11-Apr-1995
C;Accession: PS0251
R;Tsugita, A.; Kamo, M.
submitted to JIPID, April 1993
A;Reference number: PS0209
A;Accession: PS0251
A;Molecule type: protein
A;Residues: 1-15 <TSU>
A;Experimental source: germ, strain Nihonbare
C;Comment: molecular weight 15K, pI 9.2.

Query Match 51.9%; Score 27; DB 2; Length 15;
Best Local Similarity 33.3%; Pred. No. 1.1e+02;
Matches 3; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

```
Qy       1 PVVAESPKK 9
           |::|: |::
Db       5 PIMADXPRE 13
```

RESULT 2
PN0666
dystrophin-associated glycoprotein A3a-V - rabbit (fragment)
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 19-May-1994 #sequence_revision 19-May-1994 #text_change 07-May-1999
C;Accession: PN0666
R;Yoshida, M.; Mizuno, Y.; Nonaka, I.; Ozawa, E.
J. Biochem. 114, 634-639, 1993
A;Title: A dystrophin-associated glycoprotein, A3a (one of 43DAG doublets), is retair
A;Reference number: PN0662; MUID:94156881
A;Accession: PN0666
A;Molecule type: protein
A;Residues: 1-14 <YOS>
C;Comment: This protein is retained in Duchenne type muscular dystrophy muscle.
C;Keywords: glycoprotein; skeletal muscle

Query Match 50.0%; Score 26; DB 2; Length 14;
Best Local Similarity 40.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

```
Qy       1 PVVAESPKKP 10
           |:: | |  |
Db       3 PLLXEXPXAP 12
```

RESULT 3

S01810
hemoglobin AIV - tube worm (Lamellibrachia sp.) (fragment)
C;Species: Lamellibrachia sp.
C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 18-Jun-1993
C;Accession: S01810
R;Suzuki, T.; Takagi, T.; Ohta, S.
Biochem. J. 255, 541-545, 1988
A;Title: N-terminal amino acid sequence of the deep-sea tube worm haemoglobin remarkably
A;Reference number: S01807; MUID:89076216
A;Accession: S01810
A;Molecule type: protein
A;Residues: 1-32 <SUZ>

Query Match 50.0%; Score 26; DB 2; Length 32;
Best Local Similarity 83.3%; Pred. No. 3.6e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 VAESPK 8
||:||
Db 4 VAEAPK 9

RESULT 4
E47395
histone H1 II-1 (clone L92) - midge (Chironomus thummi piger) (fragment)
C;Species: Chironomus thummi piger
C;Date: 24-Feb-1994 #sequence_revision 18-Nov-1994 #text_change 03-May-1996
C;Accession: E47395
R;Schulze, E.; Trieschmann, L.; Schulze, B.; Schmidt, E.R.; Pitzel, S.; Zechel, K.; Gros
Proc. Natl. Acad. Sci. U.S.A. 90, 2481-2485, 1993
A;Title: Structural and functional differences between histone H1 sequence variants with
A;Reference number: A47395; MUID:93211985
A;Accession: E47395
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-47 <SCH>
A;Note: sequence extracted from NCBI backbone (NCBIP:128540)
C;Superfamily: histone H1

Query Match 50.0%; Score 26; DB 2; Length 47;
Best Local Similarity 50.0%; Pred. No. 5.3e+02;
Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 PVVAESPKKP 10
| |: |||
Db 10 PAKAKKEKKP 19

RESULT 5
S14064
hypothetical protein 2 - Streptomyces griseus (fragment)
C;Species: Streptomyces griseus
C;Date: 19-Mar-1997 #sequence_revision 26-Feb-1998 #text_change 26-Feb-1998
C;Accession: S14064
R;Vigal, T.; Gil, J.A.; Daza, A.; Garcia-Gonzalez, M.D.; Martin, J.F.
Mol. Gen. Genet. 225, 278-288, 1991
A;Title: Cloning, characterization and expression of an alpha-amylase gene from Streptom
A;Reference number: S14062; MUID:91172128
A;Accession: S14064
A;Molecule type: DNA
A;Residues: 1-47 <VIG>
A;Experimental source: strain IMRU3570

Query Match 50.0%; Score 26; DB 2; Length 47;
Best Local Similarity 40.0%; Pred. No. 5.3e+02;
Matches 4; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 PVVAESPKKP 10
| |::| |
Db 22 PATADTPDAP 31

RESULT 6
I53671
neurofilament heavy subunit - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 05-Nov-1999
C;Accession: I53671
R;Figlewicz, D.A.; Rouleau, G.A.; Krizus, A.; Julien, J.P.
Gene 132, 297-300, 1993
A;Title: Polymorphism in the multi-phosphorylation domain of the human neurofilament
A;Reference number: I53671; MUID:94040777
A;Accession: I53671
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-20 <RES>
A;Cross-references: GB:S66488; NID:g452861; PIDN:AAB28609.1; PID:g452862
C;Genetics:
A;Gene: GDB:NEFH
A;Cross-references: GDB:120225; OMIM:162230
A;Map position: 22q12.1-22q13.1

Query Match 48.1%; Score 25; DB 2; Length 20;
Best Local Similarity 55.6%; Pred. No. 3.4e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PVVAESPKK 9
| |:||:|
Db 2 PEKAKSPEK 10

RESULT 7
S59482
hydroxyproline-rich cell wall glycoprotein, 136K, major component - kidney bean (fra
N;Alternate names: extensin-like protein
C;Species: Phaseolus vulgaris (kidney bean)
C;Date: 27-Apr-1996 #sequence_revision 19-Jul-1996 #text_change 05-Dec-1998
C;Accession: S59482
R;Wojtaszek, P.; Trethowan, J.; Bolwell, G.P.
Plant Mol. Biol. 28, 1075-1087, 1995
A;Title: Specificity in the immobilisation of cell wall proteins in response to diff
A;Reference number: S59481; MUID:96011753
A;Accession: S59482
A;Molecule type: protein
A;Residues: 1-30 <WOJ>
C;Keywords: glycoprotein; hydroxyproline
F;8,9,10,11,12,17,18,19,20,26,27,28,29/Modified site: hydroxyproline (Pro) #status e

Query Match 48.1%; Score 25; DB 2; Length 30;
Best Local Similarity 50.0%; Pred. No. 5.1e+02;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 PVVAESPKKP 10
| | || |
Db 11 PPVVSSPPPP 20

RESULT 8
A82208
hypothetical protein VC1385 [imported] - Vibrio cholerae (strain N16961 serogroup O1
C;Species: Vibrio cholerae
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C;Accession: A82208
R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Seller
l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A;Reference number: A82035; MUID:20406833
A;Accession: A82208
A;Status: preliminary

A;Molecule type: DNA
A;Residues: 1-36 <HEI>
A;Cross-references: GB:AE004217; GB:AE003852; NID:g9655866; PIDN:AAF94543.1; GSPDB:GN001
A;Experimental source: serogroup O1; strain N16961; biotype El Tor
C;Genetics:
A;Gene: VC1385
A;Map position: 1

Query Match 48.1%; Score 25; DB 2; Length 36;
Best Local Similarity 71.4%; Pred. No. 6.2e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 AESPKKP 10
| ||| |
Db 10 APSPKAP 16

RESULT 9
R5BS36
ribosomal protein L36 - Bacillus stearothermophilus
N;Alternate names: ribosomal protein BL38; ribosomal protein II
C;Species: Bacillus stearothermophilus
C;Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 07-May-1999
C;Accession: S08566; S59066
R;Tanaka, I.; Kimura, M.; Kimura, J.; Dijk, J.
FEBS Lett. 166, 343-346, 1984
A;Title: The amino acid sequence of two small ribosomal proteins from Bacillus stearothe
A;Reference number: S07236; MUID:84108949
A;Accession: S08566
A;Molecule type: protein
A;Residues: 1-37 <TAN>
R;Urlaub, H.; Kruft, V.; Bischof, O.; Mueller, E.C.; Wittmann-Liebold, B.
EMBO J. 14, 4578-4588, 1995
A;Title: Protein-rRNA binding features and their structural and functional implications
A;Reference number: S59051; MUID:96003638
A;Accession: S59066
A;Molecule type: protein
A;Residues: 14-31 <URL>
C;Superfamily: Escherichia coli ribosomal protein L36
C;Keywords: protein biosynthesis; ribosome
F;1-37/Product: ribosomal protein L36a #status predicted <MAT>

Query Match 48.1%; Score 25; DB 1; Length 37;
Best Local Similarity 57.1%; Pred. No. 6.4e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 VVAESPK 8
|: |:||
Db 25 VICENPK 31

RESULT 10
B32307
ribosomal protein L36 - Bacillus subtilis
N;Alternate names: ribosomal protein B (rpmJ)
C;Species: Bacillus subtilis
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 20-Jun-2000
C;Accession: A32307; D69698; B32307
R;Boylan, S.A.; Suh, J.W.; Thomas, S.M.; Price, C.W.
J. Bacteriol. 171, 2553-2562, 1989
A;Title: Gene encoding the alpha core subunit of Bacillus subtilis RNA polymerase is cot
A;Reference number: A32307; MUID:89213940
A;Accession: A32307
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-37 <BOY>
A;Cross-references: GB:M26414; NID:g142458; PIDN:AAA22214.1; PID:g142460
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997

A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Ga
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo,
Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardin
A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Ma
y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portet
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scan
A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; S
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiy
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshid
A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtil
A;Reference number: A69580; MUID:98044033
A;Accession: D69698
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-37 <KUN>
A;Cross-references: GB:Z99104; GB:AL009126; NID:g2632267; PIDN:CAB11916.1; PID:g2632
A;Experimental source: strain 168
C;Genetics:
A;Gene: rpmJ
C;Superfamily: Escherichia coli ribosomal protein L36
C;Keywords: protein biosynthesis; ribosome

Query Match 48.1%; Score 25; DB 2; Length 37;
Best Local Similarity 57.1%; Pred. No. 6.4e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 VVAESPK 8
|: |:||
Db 25 VICENPK 31

RESULT 11
T44407
ribosomal protein L36 rpmJ [imported] - Bacillus halodurans (strain C-125)
C;Species: Bacillus halodurans
C;Date: 31-Jan-2000 #sequence_revision 31-Jan-2000 #text_change 15-Jun-2001
C;Accession: T44407; G83669
R;Takami, H.; Takaki, Y.; Nakasone, K.; Hirama, C.; Inoue, A.; Horikoshi, K.
Biosci. Biotechnol. Biochem. 63, 452-455, 1999
A;Title: Sequence analysis of a 32-kb region including the major ribosomal protein g
A;Reference number: Z22756; MUID:99209008
A;Accession: T44407
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-37 <TAK>
A;Cross-references: EMBL:AB017508; NID:g4512395; PIDN:BAA75295.1; PID:g4512428
A;Experimental source: strain C-125
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.;
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans
A;Reference number: A83650; MUID:20512582; PMID:11058132
A;Accession: G83669
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-37 <STO>
A;Cross-references: GB:AP001507; GB:BA000004; NID:g10172612; PIDN:BAB03878.1; GSPDB:G
A;Experimental source: strain C-125
C;Genetics:
A;Gene: rpmJ
C;Superfamily: Escherichia coli ribosomal protein L36

Query Match 48.1%; Score 25; DB 2; Length 37;
Best Local Similarity 57.1%; Pred. No. 6.4e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 VVAESPK 8
|: |:||
Db 25 VICENPK 31

RESULT 12
E97282
ribosomal protein L36 [imported] - Clostridium acetobutylicum
C;Species: Clostridium acetobutylicum
C;Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001
C;Accession: E97282
R;Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo
A;Reference number: A96900; MUID:21359325; PMID:21359325
A;Accession: E97282
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-37 <KUR>
A;Cross-references: GB:AE001437; PIDN:AAK81048.1; PID:g15026175; GSPDB:GN00168
A;Experimental source: Clostridium acetobutylicum ATCC824
C;Genetics:
A;Gene: CAC3108

Query Match 48.1%; Score 25; DB 2; Length 37;
Best Local Similarity 57.1%; Pred. No. 6.4e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 VVAESPK 8
|: |:||
Db 25 VICENPK 31

RESULT 13
F90019
50S ribosomal protein L36 [imported] - Staphylococcus aureus (strain N315)
C;Species: Staphylococcus aureus
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C;Accession: F90019
R;Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A;Title: Whole genome sequencing of meticillin-resistant Stapylococcus aureus.
A;Reference number: A89758; MUID:21311952; PMID:11418146
A;Accession: F90019
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-37 <KUR>
A;Cross-references: GB:BA000018; PID:g13702027; PIDN:BAB43319.1; GSPDB:GN00149
A;Experimental source: strain N315
C;Genetics:
A;Gene: rpmJ

Query Match 48.1%; Score 25; DB 2; Length 37;
Best Local Similarity 57.1%; Pred. No. 6.4e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 VVAESPK 8
|: |:||
Db 25 VICENPK 31

RESULT 14
AH1776
ribosomal protein L36 [imported] - Listeria innocua (strain Clip11262)
C;Species: Listeria innocua
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C;Accession: AH1776
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A;Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AH1776
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-37 <GLA>
A;Cross-references: GB:AL592022; PIDN:CAC97984.1; PID:g16415294; GSPDB:GN00178
A;Experimental source: strain Clip11262
C;Genetics:
A;Gene: rpmJ

Query Match 48.1%; Score 25; DB 2; Length 37;
Best Local Similarity 57.1%; Pred. No. 6.4e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 VVAESPK 8
|: |:||
Db 25 VICENPK 31

RESULT 15
AI1400
ribosomal protein L36 [imported] - Listeria monocytogenes (strain EGD-e)
C;Species: Listeria monocytogenes
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C;Accession: AI1400
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloe
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehl
A;Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AI1400
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-37 <GLA>
A;Cross-references: GB:NC_003210; PIDN:CAD00687.1; PID:g16412097; GSPDB:GN00177
A;Experimental source: strain EGD-e
C;Genetics:
A;Gene: rpmJ

Query Match 48.1%; Score 25; DB 2; Length 37;
Best Local Similarity 57.1%; Pred. No. 6.4e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 VVAESPK 8
|: |:||
Db 25 VICENPK 31

Search completed: October 12, 2002, 20:49:44
Job time: 248 sec

GenCore version 4.5

OM protein - protein search, using sw model

Run on: October 12, 2002, 20:48:06 ; Search time 9.8 Seconds
(without alignments)
39.510 Million cell updates/sec

Title: US-09-408-578A-1
Perfect score: 52
Sequence: 1 PVVAESPKKP 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 3667

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	26	50.0	32	1	GLB4_LAMSP	P20413 lamellibrac
2	25	48.1	29	1	TAT_HV1Z3	P12510 human immun
3	25	48.1	33	1	RL4_HALCU	P05967 halobacteri
4	25	48.1	37	1	RL36_BACHD	O50631 bacillus ha
5	25	48.1	37	1	RL36_BACST	P07841 bacillus st
6	25	48.1	37	1	RL36_BACSU	P20278 bacillus su
7	25	48.1	37	1	RL36_THETH	P80256 thermus aqu
8	25	48.1	39	1	PSAX_ANAVA	P23319 anabaena va
9	25	48.1	49	1	YKQ6_CAEEL	P34301 caenorhabdi
10	25	48.1	50	1	YKQ2_CAEEL	P34297 caenorhabdi
11	25	48.1	50	1	YKQ5_CAEEL	P34300 caenorhabdi
12	24	46.2	16	1	H5_COTJA	P18638 coturnix co
13	24	46.2	37	1	CHCD_ANTPO	P08931 antheraea p
14	24	46.2	37	1	RL36_CAMJE	Q9pm84 campylobact
15	24	46.2	40	1	H32_TETBO	P17319 tetrahymena
16	24	46.2	45	1	PHI3_MYTCA	P11860 mytilus cal
17	24	46.2	50	1	HOKE_ECOLI	P77091 escherichia
18	23	44.2	17	1	TRP2_LEUMA	P81733 leucophaea
19	23	44.2	24	1	CR16_LITXA	P56231 litoria xan
20	23	44.2	24	1	CR17_LITXA	P56232 litoria xan
21	23	44.2	24	1	CR18_LITCH	P81251 litoria chl
22	23	44.2	24	1	CR19_LITCH	P81252 litoria chl
23	23	44.2	25	1	CR11_LITSP	P56226 litoria spl
24	23	44.2	25	1	CR12_LITCE	P56227 litoria cae
25	23	44.2	25	1	CR13_LITCE	P56228 litoria cae
26	23	44.2	25	1	CR14_LITGI	P56229 litoria gil
27	23	44.2	25	1	CR15_LITCE	P56230 litoria cae
28	23	44.2	25	1	CR1A_LITSP	P82104 litoria spl
29	23	44.2	36	1	Y297_ARCFU	O29945 archaeoglob
30	23	44.2	39	1	GVPC_SPICC	P81000 spirulina s
31	23	44.2	44	1	BAB4_BABBO	P14201 babesia bov
32	22	42.3	17	1	APID_BOMPA	P81464 bombus pasc
33	22	42.3	34	1	H1S_STRPU	P19376 strongyloce
34	22	42.3	37	1	RK36_CHLVU	P56360 chlorella v
35	22	42.3	41	1	LAMA_EMENI	P38095 emericella
36	21	40.4	17	1	SRY_URSAR	P36396 ursus arcto
37	21	40.4	20	1	HELT_HELHO	P46693 heloderma h
38	21	40.4	20	1	MI17_BOVIN	P35451 bos taurus
39	21	40.4	31	1	CU54_LOCMI	P11738 locusta mig
40	21	40.4	33	1	PSBT_MAIZE	P37257 zea mays (m
41	21	40.4	34	1	EM1_ENSMI	P27205 ensis minor
42	21	40.4	34	1	PSBT_TOBAC	P12184 nicotiana t
43	21	40.4	35	1	COPA_CANFA	P40765 canis famil
44	21	40.4	35	1	PSBT_MARPO	P12182 marchantia
45	21	40.4	35	1	PSBT_ORYSA	P12183 oryza sativ

ALIGNMENTS

RESULT 1
GLB4_LAMSP
ID GLB4_LAMSP STANDARD; PRT; 32 AA.
AC P20413;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-FEB-1991 (Rel. 17, Last annotation update)
DE Giant hemoglobin AIV chain (Fragment).
OS Lamellibrachia sp. (Deep-sea giant tube worm).
OC Eukaryota; Metazoa; Vestimentifera; Basibranchia; Lamellibrachiida;
OC Lamellibrachiidae; Lamellibrachia.
OX NCBI_TaxID=6424;
RN [1]
RP SEQUENCE.
RX MEDLINE=89076216; PubMed=3202832;
RA Suzuki T., Takagi T., Ohta S.;
RT "N-terminal amino acid sequence of the deep-sea tube worm haemoglobin
RT remarkably resembles that of annelid haemoglobin.";
RL Biochem. J. 255:541-545(1988).
CC -!- SUBUNIT: GIANT HEMOGLOBIN IS COMPOSED OF FOUR HEME-CONTAINING
CC CHAINS (AI TO AIV), AND TWO LINKER CHAINS (AV AND AVI).
DR PIR; S01810; S01810.
DR InterPro; IPR000971; Globin.
DR PROSITE; PS01033; GLOBIN; PARTIAL.
KW Heme; Oxygen transport; Transport.
FT NON_TER 32 32
SQ SEQUENCE 32 AA; 3695 MW; FD7E1ADABD35FD13 CRC64;

Query Match 50.0%; Score 26; DB 1; Length 32;
Best Local Similarity 83.3%; Pred. No. 1.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy      3 VAESPK 8
          |||:||
Db      4 VAEAPK 9
```

RESULT 2
TAT_HV1Z3
ID TAT_HV1Z3 STANDARD; PRT; 29 AA.
AC P12510;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE TAT protein (Transactivating regulatory protein) (Fragment).
GN TAT.
OS Human immunodeficiency virus type 1 (Zaire 3 isolate) (HIV-1).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11680;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86259728; PubMed=3014529;
RA Willey R.W., Rutledge R.A., Dias S., Folks T., Theodore T.,
RA Buckler C.E., Martin M.A.;
RT "Identification of conserved and divergent domains within the

```
RT   envelope gene of the acquired immunodeficiency syndrome retrovirus.";
RL   Proc. Natl. Acad. Sci. U.S.A. 83:5038-5042(1986).
CC   -!- FUNCTION: TRANSCRIPTIONAL REGULATOR THAT ACTS BY BINDING TO THE
CC       TRANS-ACTIVATING RESPONSIVE SEQUENCE (TAR) RNA ELEMENT AND
CC       ACTIVATES TRANSCRIPTION INITIATION AND/OR ELONGATION FROM THE LTR
CC       PROMOTER.
CC   -!- SUBUNIT: BINDS CYCLIN T1 (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLAR.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; K03347; AAA45374.1; -.
DR   HIV; K03347; TAT$Z3.
KW   Transcription regulation; Activator; RNA-binding; Nuclear protein;
KW   AIDS.
FT   NON_TER       1      1
SQ   SEQUENCE   29 AA;  3229 MW;  A8C8E45E70F2C192 CRC64;


  Query Match             48.1%;  Score 25;  DB 1;  Length 29;
  Best Local Similarity   55.6%;  Pred. No. 2.6e+02;
  Matches    5;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 PVVAESPKK 9
            |   | |||
Db       9 PTGQEEPKK 17


RESULT   3
RL4_HALCU
ID   RL4_HALCU      STANDARD;      PRT;    33 AA.
AC   P05967;
DT   01-NOV-1988 (Rel. 09, Created)
DT   01-NOV-1988 (Rel. 09, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   50S ribosomal protein L4E (HL9) (Fragment).
GN   RPL4E.
OS   Halobacterium cutirubrum.
OC   Archaea; Euryarchaeota; Halobacteriales; Halobacteriaceae;
OC   Halobacterium.
OX   NCBI_TaxID=2242;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=84282108; PubMed=6467081;
RA   Matheson A.T., Yaguchi M., Christensen P., Rollin C.F., Hasnain S.;
RT   "Purification, properties, and N-terminal amino acid sequence of
RT   certain 50S ribosomal subunit proteins from the archaebacterium
RT   Halobacterium cutirubrum.";
RL   Can. J. Biochem. Cell Biol. 62:426-433(1984).
CC   -!- SIMILARITY: BELONGS TO THE L4E FAMILY OF RIBOSOMAL PROTEINS.
DR   PIR; S08551; S08551.
DR   InterPro; IPR002136; Ribosomal_L4/L1E.
DR   PROSITE; PS00939; RIBOSOMAL_L1E; PARTIAL.
KW   Ribosomal protein.
FT   NON_TER      33     33
SQ   SEQUENCE   33 AA;  3753 MW;  E0D799DAAF1E3B7B CRC64;


  Query Match             48.1%;  Score 25;  DB 1;  Length 33;
  Best Local Similarity   50.0%;  Pred. No. 2.9e+02;
  Matches    5;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 PVVAESPKKP 10
            | | | | :|
Db      19 PRVFEEPVRP 28


RESULT   4
RL36_BACHD
ID   RL36_BACHD     STANDARD;      PRT;    37 AA.
AC   O50631; Q9JPW6;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   50S ribosomal protein L36.
GN   RPMJ OR BH0159.
OS   Bacillus halodurans.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group;
OC   Bacillus/Staphylococcus group; Bacillus.
OX   NCBI_TaxID=86665;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C-125 / JCM 9153;
RX   MEDLINE=99052103; PubMed=9835038;
RA   Nakasone K., Takaki Y., Takami H., Inoue A., Horikoshi K.;
RT   "Cloning and expression of the gene encoding RNA polymerase alpha
RT   subunit from alkaliphilic Bacillus sp. strain C-125.";
RL   FEMS Microbiol. Lett. 168:269-276(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C-125 / JCM 9153;
RX   MEDLINE=99209008; PubMed=10192928;
RA   Takami H., Takaki Y., Nakasone K., Hirama C., Inoue A., Horikoshi K.;
RT   "Sequence analysis of a 32-kb region including the major ribosomal
RT   protein gene clusters from alkaliphilic Bacillus sp. strain C-125.";
RL   Biosci. Biotechnol. Biochem. 63:452-455(1999).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C-125 / JCM 9153;
RX   MEDLINE=20512582; PubMed=11058132;
RA   Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA   Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA   Horikoshi K.;
RT   "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT   halodurans and genomic sequence comparison with Bacillus subtilis.";
RL   Nucleic Acids Res. 28:4317-4331(2000).
CC   -!- SIMILARITY: BELONGS TO THE L36P FAMILY OF RIBOSOMAL PROTEINS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AB010082; BAA24191.1; -.
DR   EMBL; AB017508; BAA75295.1; -.
DR   EMBL; AP001507; BAB03878.1; -.
DR   HSSP; P80256; 1DFE.
DR   InterPro; IPR000473; Ribosomal_L36.
DR   Pfam; PF00444; Ribosomal_L36; 1.
DR   ProDom; PD002101; Ribosomal_L36; 1.
DR   PROSITE; PS00828; RIBOSOMAL_L36; 1.
KW   Ribosomal protein; Complete proteome.
SQ   SEQUENCE   37 AA;  4278 MW;  93A9A9E0714F4CEF CRC64;


  Query Match             48.1%;  Score 25;  DB 1;  Length 37;
  Best Local Similarity   57.1%;  Pred. No. 3.2e+02;
  Matches    4;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 VVAESPK 8
            |: |:||
Db      25 VICENPK 31


RESULT   5
RL36_BACST
ID   RL36_BACST     STANDARD;      PRT;    37 AA.
```

```
AC   P07841;
DT   01-AUG-1988 (Rel. 08, Created)
DT   01-AUG-1988 (Rel. 08, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   50S ribosomal protein L36 (Ribosomal protein II) (Ribosomal protein B)
DE   (BL38).
GN   RPMJ.
OS   Bacillus stearothermophilus.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group;
OC   Bacillus/Staphylococcus group; Geobacillus.
OX   NCBI_TaxID=1422;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=84108949; PubMed=6420194;
RA   Tanaka I., Kimura M., Kimura J., Dijk J.;
RT   "The amino acid sequence of two small ribosomal proteins from
RT   Bacillus stearothermophilus.";
RL   FEBS Lett. 166:343-346(1984).
CC   -!- SIMILARITY: BELONGS TO THE L36P FAMILY OF RIBOSOMAL PROTEINS.
DR   PIR; S08566; R5BS36.
DR   HSSP; P80256; 1DFE.
DR   InterPro; IPR000473; Ribosomal_L36.
DR   Pfam; PF00444; Ribosomal_L36; 1.
DR   ProDom; PD002101; Ribosomal_L36; 1.
DR   PROSITE; PS00828; RIBOSOMAL_L36; 1.
KW   Ribosomal protein.
SQ   SEQUENCE   37 AA;  4361 MW;  7F6B0920714F4CF7 CRC64;


  Query Match             48.1%;  Score 25;  DB 1;  Length 37;
  Best Local Similarity   57.1%;  Pred. No. 3.2e+02;
  Matches    4;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 VVAESPK 8
           |: |:||
Db      25 VICENPK 31


RESULT   6
RL36_BACSU
ID   RL36_BACSU     STANDARD;      PRT;    37 AA.
AC   P20278;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   50S ribosomal protein L36 (Ribosomal protein II) (Ribosomal protein B)
DE   (BL38).
GN   RPMJ.
OS   Bacillus subtilis.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group;
OC   Bacillus/Staphylococcus group; Bacillus.
OX   NCBI_TaxID=1423;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89213940; PubMed=2496109;
RA   Boylan S.A., Suh J.-W., Thomas S.M., Price C.W.;
RT   "Gene encoding the alpha core subunit of Bacillus subtilis RNA
RT   polymerase is cotranscribed with the genes for initiation factor 1
RT   and ribosomal proteins B, S13, S11, and L17.";
RL   J. Bacteriol. 171:2553-2562(1989).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=168 / MARBURG;
RX   MEDLINE=96186897; PubMed=8635744;
RA   Suh J.W., Boylan S.A., Oh S.H., Price C.W.;
RT   "Genetic and transcriptional organization of the Bacillus subtilis
RT   spc-alpha region.";
RL   Gene 169:17-23(1996).
CC   -!- SIMILARITY: BELONGS TO THE L36P FAMILY OF RIBOSOMAL PROTEINS.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; M26414; AAA22214.1; -.
DR   EMBL; L47971; AAB06823.1; -.
DR   EMBL; Z99104; CAB11916.1; -.
DR   PIR; B32307; B32307.
DR   HSSP; P80256; 1DFE.
DR   SubtiList; BG11042; rpmJ.
DR   InterPro; IPR000473; Ribosomal_L36.
DR   Pfam; PF00444; Ribosomal_L36; 1.
DR   ProDom; PD002101; Ribosomal_L36; 1.
DR   PROSITE; PS00828; RIBOSOMAL_L36; 1.
KW   Ribosomal protein; Complete proteome.
SQ   SEQUENCE   37 AA;  4305 MW;  7F79A9E0714F4CF7 CRC64;


  Query Match             48.1%;  Score 25;  DB 1;  Length 37;
  Best Local Similarity   57.1%;  Pred. No. 3.2e+02;
  Matches    4;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 VVAESPK 8
           |: |:||
Db      25 VICENPK 31


RESULT   7
RL36_THETH
ID   RL36_THETH     STANDARD;      PRT;    37 AA.
AC   P80256;
DT   01-OCT-1993 (Rel. 27, Created)
DT   01-OCT-1993 (Rel. 27, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   50S ribosomal protein L36 (Ribosomal protein B).
GN   RPMJ OR RPL36.
OS   Thermus aquaticus (subsp. thermophilus).
OC   Bacteria; Thermus/Deinococcus group; Thermus group; Thermus.
OX   NCBI_TaxID=274;
RN   [1]
RP   SEQUENCE.
RC   STRAIN=HB8 / ATCC 27634;
RA   Boysen R.I., Schroeder W., Erdmann V.A.;
RL   Submitted (SEP-1993) to the SWISS-PROT data bank.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=HB8 / ATCC 27634;
RX   MEDLINE=99098837; PubMed=9880810;
RA   Wada T., Yamazaki T., Kuramitsu S., Kyogoku Y.;
RT   "Cloning of the RNA polymerase alpha subunit gene from Thermus
RT   thermophilus HB8 and characterization of the protein.";
RL   J. Biochem. 125:143-150(1999).
RN   [3]
RP   STRUCTURE BY NMR.
RX   MEDLINE=20124006; PubMed=10656825;
RA   Hard T., Rak A., Allard P., Kloo L., Garber M.;
RT   "The solution structure of ribosomal protein L36 from Thermus
RT   thermophilus reveals a zinc-ribbon-like fold.";
RL   J. Mol. Biol. 296:169-180(2000).
CC   -!- SIMILARITY: BELONGS TO THE L36P FAMILY OF RIBOSOMAL PROTEINS.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AB024328; BAA75545.1; -.
DR   PDB; 1DFE; 16-FEB-00.
DR   PDB; 1DGZ; 08-DEC-99.
```

```
DR   InterPro; IPR000473; Ribosomal_L36.
DR   Pfam; PF00444; Ribosomal_L36; 1.
DR   ProDom; PD002101; Ribosomal_L36; 1.
DR   PROSITE; PS00828; RIBOSOMAL_L36; 1.
KW   Ribosomal protein; 3D-structure.
SQ   SEQUENCE   37 AA;  4421 MW;  439072E1737E2AE8 CRC64;


  Query Match             48.1%;  Score 25;  DB 1;  Length 37;
  Best Local Similarity   57.1%;  Pred. No. 3.2e+02;
  Matches    4;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 VVAESPK 8
           |: |:||
Db      25 VICENPK 31


RESULT   8
PSAX_ANAVA
ID   PSAX_ANAVA     STANDARD;      PRT;    39 AA.
AC   P23319;
DT   01-NOV-1991 (Rel. 20, Created)
DT   01-NOV-1991 (Rel. 20, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Photosystem I 4.8 kDa protein (Fragment).
GN   PSAX.
OS   Anabaena variabilis.
OC   Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Anabaena.
OX   NCBI_TaxID=1172;
RN   [1]
RP   SEQUENCE.
RC   STRAIN=PCC 7937 / ATCC 29413;
RX   MEDLINE=91348228; PubMed=1908790;
RA   Ikeuchi M., Nyhus K.J., Inoue Y., Pakrasi H.B.;
RT   "Identities of four low-molecular-mass subunits of the photosystem I
RT   complex from Anabaena variabilis ATCC 29413. Evidence for the
RT   presence of the psaI gene product in a cyanobacterial complex.";
RL   FEBS Lett. 287:5-9(1991).
CC   -!- SIMILARITY: BELONGS TO THE PSAX FAMILY.
DR   PIR; S16978; S16978.
KW   Photosystem I; Photosynthesis.
FT   NON_TER      39      39
SQ   SEQUENCE   39 AA;  4141 MW;  E762CF2381CDFCA7 CRC64;


  Query Match             48.1%;  Score 25;  DB 1;  Length 39;
  Best Local Similarity   50.0%;  Pred. No. 3.4e+02;
  Matches    5;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 PVVAESPKKP 10
           | || :  ||
Db       6 PAVANTGAKP 15


RESULT   9
YKQ6_CAEEL
ID   YKQ6_CAEEL     STANDARD;      PRT;    49 AA.
AC   P34301;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Hypothetical 5.0 kDa protein C06E1.6 in chromosome III.
GN   C06E1.6.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RX   MEDLINE=94150718; PubMed=7906398;
RA   Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA   Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA   Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
RA   Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
RA   Johnston L., Jones M., Kershaw J., Kirsten J., Laisster N.,
RA   Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
RA   Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,
RA   Sims M., Smaldon N., Smith A., Smith M., Sonnhammer E., Staden R.,
RA   Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
RA   Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J.,
RA   Wohldman P.;
RT   "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT   elegans.";
RL   Nature 368:32-38(1994).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L16559; AAA27937.1; -.
DR   WormPep; C06E1.6; CE00061.
KW   Hypothetical protein.
SQ   SEQUENCE   49 AA;  4970 MW;  8FB975FB9E1AA49B CRC64;


  Query Match             48.1%;  Score 25;  DB 1;  Length 49;
  Best Local Similarity   83.3%;  Pred. No. 4.2e+02;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       4 AESPKK 9
           || |||
Db      44 AERPKK 49


RESULT  10
YKQ2_CAEEL
ID   YKQ2_CAEEL     STANDARD;      PRT;    50 AA.
AC   P34297;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Hypothetical 5.0 kDa protein C06E1.2 in chromosome III.
GN   C06E1.2.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RX   MEDLINE=94150718; PubMed=7906398;
RA   Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA   Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA   Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
RA   Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
RA   Johnston L., Jones M., Kershaw J., Kirsten J., Laisster N.,
RA   Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
RA   Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,
RA   Sims M., Smaldon N., Smith A., Smith M., Sonnhammer E., Staden R.,
RA   Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
RA   Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J.,
RA   Wohldman P.;
RT   "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT   elegans.";
RL   Nature 368:32-38(1994).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
```

```
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; L16559; AAA27936.1; -.
DR   WormPep; C06E1.2; CE00057.
KW   Hypothetical protein.
SQ   SEQUENCE   50 AA;  5038 MW;  F28BBF1F40059EB7 CRC64;


  Query Match             48.1%;  Score 25;  DB 1;  Length 50;
  Best Local Similarity   83.3%;  Pred. No. 4.3e+02;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      4 AESPKK 9
          || |||
Db     45 AERPKK 50


RESULT  11
YKQ5_CAEEL
ID   YKQ5_CAEEL     STANDARD;      PRT;    50 AA.
AC   P34300;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Hypothetical 5.0 kDa protein C06E1.5 in chromosome III.
GN   C06E1.5.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RX   MEDLINE=94150718; PubMed=7906398;
RA   Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA   Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA   Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
RA   Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
RA   Johnston L., Jones M., Kershaw J., Kirsten J., Laisster N.,
RA   Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
RA   Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,
RA   Sims M., Smaldon N., Smith A., Smith M., Sonnhammer E., Staden R.,
RA   Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
RA   Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J.,
RA   Wohldman P.;
RT   "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT   elegans.";
RL   Nature 368:32-38(1994).
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; L16559; AAA27935.1; -.
DR   WormPep; C06E1.5; CE00060.
KW   Hypothetical protein.
SQ   SEQUENCE   50 AA;  4954 MW;  B6D641A991D520FB CRC64;


  Query Match             48.1%;  Score 25;  DB 1;  Length 50;
  Best Local Similarity   83.3%;  Pred. No. 4.3e+02;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      4 AESPKK 9
          || |||
Db     45 AERPKK 50


RESULT  12
H5_COTJA
ID   H5_COTJA       STANDARD;      PRT;    16 AA.
AC   P18638;
DT   01-NOV-1990 (Rel. 16, Created)
DT   01-NOV-1990 (Rel. 16, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Histone H5 (Fragment).
OS   Coturnix coturnix japonica (Japanese quail).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Coturnix.
OX   NCBI_TaxID=93934;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=76277939; PubMed=962913;
RA   Seligy V., Roy C., Dove M., Yaguchi M.;
RT   "Species variability of N-terminal sequence of avian erythrocyte-
RT   specific histone H5.";
RL   Biochem. Biophys. Res. Commun. 71:196-202(1976).
CC   -!- FUNCTION: HISTONE H5 PERFORMS THE SAME FUNCTION AS H1, BEING
CC       NECESSARY FOR THE CONDENSATION OF NUCLEOSOME CHAINS INTO HIGHER
CC       ORDER STRUCTURES, AND REPLACES HISTONE H1 IN CERTAIN CELLS.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- TISSUE SPECIFICITY: ERYTHROID CELLS.
CC   -!- SIMILARITY: BELONGS TO THE HISTONE H1/H5 FAMILY.
KW   Chromosomal protein; Nuclear protein; DNA-binding; DNA condensation.
FT   NON_TER      16     16
SQ   SEQUENCE   16 AA;  1665 MW;  DB528219B3074D3C CRC64;


  Query Match             46.2%;  Score 24;  DB 1;  Length 16;
  Best Local Similarity   44.4%;  Pred. No. 2.2e+02;
  Matches    4;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      2 VVAESPKKP 10
          |:: :| ||
Db      5 VLSPAPAKP 13


RESULT  13
CHCD_ANTPO
ID   CHCD_ANTPO     STANDARD;      PRT;    37 AA.
AC   P08931;
DT   01-NOV-1988 (Rel. 09, Created)
DT   01-NOV-1988 (Rel. 09, Last sequence update)
DT   01-NOV-1988 (Rel. 09, Last annotation update)
DE   Chorion class CB protein PCH12 (Fragment).
OS   Antheraea polyphemus (Polyphemus moth).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC   Bombycoidea; Saturniidae; Saturniinae; Antheraea.
OX   NCBI_TaxID=7120;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=83195030; PubMed=6573656;
RA   Regier J.C., Kafatos F.C., Hamodrakas S.J.;
RT   "Silkmoth chorion multigene families constitute a superfamily:
RT   comparison of C and B family sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 80:1043-1047(1983).
CC   -!- FUNCTION: THIS PROTEIN IS ONE OF MANY FROM THE EGGSHELL OF THE
CC       SILK MOTH.
CC   -!- SIMILARITY: MEMBER OF THE BETA-BRANCH OF CHORION PROTEIN TO WHICH
CC       BELONG CLASSES B, CB AND HCB.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
```

```
DR   EMBL; V00080; CAA23422.1; -.
KW   Eggshell; Chorion; Repeat; Multigene family.
FT   NON_TER       1      1
FT   DOMAIN       <1     26       CENTRAL DOMAIN.
FT   DOMAIN       27     >37      RIGHT ARM.
FT   NON_TER      37     37
SQ   SEQUENCE   37 AA;  3615 MW;  2EF1FF446F4D532C CRC64;


  Query Match             46.2%;  Score 24;  DB 1;  Length 37;
  Best Local Similarity   66.7%;  Pred. No. 4.8e+02;
  Matches    4;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 VVAESP 7
           :|||:|
Db      23 IVAETP 28


RESULT  14
RL36_CAMJE
ID   RL36_CAMJE     STANDARD;      PRT;    37 AA.
AC   Q9PM84;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   50S ribosomal protein L36.
GN   RPMJ OR CJ1591.
OS   Campylobacter jejuni.
OC   Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;
OC   Campylobacter.
OX   NCBI_TaxID=197;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NCTC 11168;
RX   MEDLINE=20150912; PubMed=10688204;
RA   Parkhill J., Wren B.W., Mungall K., Ketley J.M., Churcher C.,
RA   Basham D., Chillingworth T., Davies R.M., Feltwell T., Holroyd S.,
RA   Jagels K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.W.,
RA   Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M.,
RA   Whitehead S., Barrell B.G.;
RT   "The genome sequence of the food-borne pathogen Campylobacter jejuni
RT   reveals hypervariable sequences.";
RL   Nature 403:665-668(2000).
CC   -!- SIMILARITY: BELONGS TO THE L36P FAMILY OF RIBOSOMAL PROTEINS.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AL139079; CAB73579.1; -.
DR   InterPro; IPR000473; Ribosomal_L36.
DR   Pfam; PF00444; Ribosomal_L36; 1.
DR   PROSITE; PS00828; RIBOSOMAL_L36; 1.
KW   Ribosomal protein; Complete proteome.
SQ   SEQUENCE   37 AA;  4364 MW;  4548AEC5256D94ED CRC64;


  Query Match             46.2%;  Score 24;  DB 1;  Length 37;
  Best Local Similarity   42.9%;  Pred. No. 4.8e+02;
  Matches    3;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 VVAESPK 8
           :: |:||
Db      25 IICENPK 31


RESULT  15
H32_TETBO
ID   H32_TETBO      STANDARD;      PRT;    40 AA.
AC   P17319;
DT   01-AUG-1990 (Rel. 15, Created)
DT   01-AUG-1990 (Rel. 15, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Histone H3.2 (Fragment).
OS   Tetrahymena borealis.
OC   Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Hymenostomatida;
OC   Tetrahymenina; Tetrahymena.
OX   NCBI_TaxID=5893;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=WZ3;
RX   MEDLINE=90221813; PubMed=2129549;
RA   Brunk C.F., Sadler L.A.;
RT   "Characterization of the promoter region of Tetrahymena genes.";
RL   Nucleic Acids Res. 18:323-329(1990).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=WZ3;
RX   MEDLINE=90219078; PubMed=2129541;
RA   Brunk C.F., Kahn R.W., Sadler L.A.;
RT   "Phylogenetic relationships among Tetrahymena species determined
RT   using the polymerase chain reaction.";
RL   J. Mol. Evol. 30:290-297(1990).
CC   -!- FUNCTION: HISTONE H3, ALONG WITH HISTONE H4, PLAYS A CENTRAL ROLE
CC       IN NUCLEOSOME FORMATION.
CC   -!- SUBUNIT: THE NUCLEOSOME IS AN OCTAMER CONTAINING TWO MOLECULES OF
CC       H2A, H2B, H3, AND H4; WHICH WRAP APPROXIMATIVELY 146 BP OF DNA.
CC   -!- SIMILARITY: BELONGS TO THE HISTONE H3 FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X17128; CAA34990.1; ALT_SEQ.
DR   PIR; S10263; S10263.
DR   InterPro; IPR000164; Histone_H3.
DR   PROSITE; PS00322; HISTONE_H3_1; PARTIAL.
DR   PROSITE; PS00959; HISTONE_H3_2; PARTIAL.
KW   Nuclear protein; Chromosomal protein; DNA-binding; Nucleosome core.
FT   INIT_MET      0      0
FT   NON_TER      40     40
SQ   SEQUENCE   40 AA;  4183 MW;  E79CED7EBE66EE02 CRC64;


  Query Match             46.2%;  Score 24;  DB 1;  Length 40;
  Best Local Similarity   66.7%;  Pred. No. 5.2e+02;
  Matches    4;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       4 AESPKK 9
           ||:|:|
Db      13 AEAPRK 18


Search completed: October 12, 2002, 20:50:45
Job time: 159 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: October 12, 2002, 20:47:16 ; Search time 25.85 Seconds
(without alignments)
66.923 Million cell updates/sec

Title: US-09-408-578A-1
Perfect score: 52
Sequence: 1 PVVAESPKKP 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 29986

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_19:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	30	57.7	22	13	Q9PS65	Q9ps65 oncorhynchu
2	28	53.8	12	4	Q9UC37	Q9uc37 homo sapien
3	28	53.8	41	9	Q9MC40	Q9mc40 bacteriopha
4	28	53.8	47	16	Q92P30	Q92p30 rhizobium m
5	27	51.9	25	7	O19452	O19452 mus musculu
6	27	51.9	37	11	O88771	O88771 rattus norv
7	27	51.9	39	2	P82570	P82570 streptococc
8	27	51.9	42	5	Q95R01	Q95r01 caenorhabdi
9	27	51.9	47	11	Q9JIC8	Q9jic8 mus musculu
10	26	50.0	29	8	Q9GB18	Q9gb18 melidectes
11	26	50.0	47	12	Q84028	Q84028 influenza a
12	25	48.1	18	4	Q96L04	Q96l04 homo sapien
13	25	48.1	20	4	Q16070	Q16070 homo sapien
14	25	48.1	22	10	Q9SLV6	Q9slv6 nicotiana t
15	25	48.1	27	4	Q9UMH7	Q9umh7 homo sapien
16	25	48.1	29	5	Q95SE3	Q95se3 drosophila
17	25	48.1	30	3	Q9UR69	Q9ur69 trametes ve
18	25	48.1	32	6	Q9TRQ1	Q9trq1 bos taurus
19	25	48.1	34	11	Q9CRT3	Q9crt3 mus musculu
20	25	48.1	36	16	Q9KS74	Q9ks74 vibrio chol
21	25	48.1	37	16	Q97EK2	Q97ek2 clostridium
22	25	48.1	37	16	Q99S42	Q99s42 staphylococ
23	25	48.1	37	16	Q927N0	Q927n0 listeria in
24	25	48.1	43	4	Q14909	Q14909 homo sapien
25	25	48.1	45	5	Q9VB16	Q9vb16 drosophila
26	25	48.1	48	2	Q9KJK4	Q9kjk4 fischerella
27	25	48.1	49	4	Q96GU7	Q96gu7 homo sapien
28	24.5	47.1	44	4	O95537	O95537 homo sapien
29	24	46.2	14	10	Q94IT6	Q94it6 fragaria nu
30	24	46.2	14	11	Q9QVF3	Q9qvf3 rattus sp.
31	24	46.2	15	10	Q9S8F1	Q9s8f1 zea mays (m
32	24	46.2	18	8	Q9ZY79	Q9zy79 idris sp. c
33	24	46.2	20	6	Q9TR50	Q9tr50 bos taurus
34	24	46.2	21	2	Q9X3K2	Q9x3k2 prochloroco
35	24	46.2	24	6	Q9BGI8	Q9bgi8 oryctolagus
36	24	46.2	25	6	Q9BGI9	Q9bgi9 cheirogaleu
37	24	46.2	28	6	Q9N0N9	Q9n0n9 cercopithec
38	24	46.2	32	4	O75210	O75210 homo sapien
39	24	46.2	32	4	Q9UPL5	Q9upl5 homo sapien
40	24	46.2	36	5	O61188	O61188 colpidium c
41	24	46.2	36	6	Q9TQR7	Q9tqr7 equus cabal
42	24	46.2	37	4	Q15685	Q15685 homo sapien
43	24	46.2	37	16	Q97T62	Q97t62 streptococc
44	24	46.2	38	2	Q9X3T8	Q9x3t8 salmonella
45	24	46.2	38	2	Q9X3U4	Q9x3u4 salmonella

ALIGNMENTS

RESULT 1
Q9PS65
ID Q9PS65 PRELIMINARY; PRT; 22 AA.
AC Q9PS65;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE HISTONE H1B.
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxID=8022;
RN [1]
RP SEQUENCE.
RX MEDLINE=92082492; PubMed=1747124;
RA Davie J.R., Delcuve G.P.;
RT "Characterization and chromatin distribution of the H1 histones and
RT high-mobility-group non-histone chromosomal proteins of trout liver
RT and hepatocellular carcinoma.";
RL Biochem. J. 280:491-497(1991).
SQ SEQUENCE 22 AA; 2132 MW; 3E90388F68189AE3 CRC64;

Query Match 57.7%; Score 30; DB 13; Length 22;
Best Local Similarity 71.4%; Pred. No. 97;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```
Qy       4 AESPKKP 10
           |::||||
Db      14 AKAPKKP 20
```

RESULT 2
Q9UC37
ID Q9UC37 PRELIMINARY; PRT; 12 AA.
AC Q9UC37;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

```
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   ALPHA B CRYSTALLIN FRAGMENT 5.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=92218434; PubMed=1560006;
RA   Kato K., Shinohara H., Goto S., Inaguma Y., Morishita R., Asano T.;
RT   "Copurification of small heat shock protein with alpha B crystallin
RT   from human skeletal muscle.";
RL   J. Biol. Chem. 267:7718-7725(1992).
SQ   SEQUENCE   12 AA;  1268 MW;  D37BD529CCC1B2CD CRC64;


  Query Match             53.8%;  Score 28;  DB 4;  Length 12;
  Best Local Similarity   55.6%;  Pred. No. 1.2e+02;
  Matches    5;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 PVVAESPKK 9
            | |  :|||
Db       4 PAVTAAPKK 12


RESULT   3
Q9MC40
ID   Q9MC40      PRELIMINARY;      PRT;    41 AA.
AC   Q9MC40;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   ORF92.
GN   ORF92.
OS   Bacteriophage D3.
OC   Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae;
OC   Lambda phage group.
OX   NCBI_TaxID=31535;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20042341; PubMed=10572124;
RA   Gilakjan Z.A., Kropinski A.M.;
RT   "Cloning and analysis of the capsid morphogenesis genes of Pseudomonas
RT   aeruginosa bacteriophage D3: another example of protein chain mail?";
RL   J. Bacteriol. 181:7221-7227(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20485557; PubMed=11029426;
RA   Kropinski A.M.;
RT   "Sequence of the Genome of the Temperate, Serotype-Converting,
RT   Pseudomonas aeruginosa Bacteriophage D3.";
RL   J. Bacteriol. 182:6066-6074(2000).
DR   EMBL; AF165214; AAF80768.1; -.
SQ   SEQUENCE   41 AA;  4629 MW;  9632B19CBD142821 CRC64;


  Query Match             53.8%;  Score 28;  DB 9;  Length 41;
  Best Local Similarity   62.5%;  Pred. No. 4.1e+02;
  Matches    5;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 VAESPKKP 10
            ||  ||:|
Db      29 VATKPKRP 36


RESULT   4
Q92P30
ID   Q92P30      PRELIMINARY;      PRT;    47 AA.
AC   Q92P30;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HYPOTHETICAL PROTEIN SMC04310.
GN   SMC04310.
OS   Rhizobium meliloti (Sinorhizobium meliloti).
OC   Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC   Rhizobiaceae; Sinorhizobium.
OX   NCBI_TaxID=382;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=1021;
RX   MEDLINE=21368234; PubMed=11474104;
RA   Galibert F., Finan T.M., Long S.R., Puehler A., Abola P., Ampe F.,
RA   Barloy-Hubler F., Barnett M.J., Becker A., Boistard P., Bothe G.,
RA   Boutry M., Bowser L., Buhrmester J., Cadieu E., Capela D., Chain P.,
RA   Cowie A., Davis R.W., Dreano S., Federspiel N.A., Fisher R.F.,
RA   Gloux S., Godrie T., Goffeau A., Golding B., Gouzy J., Gurjal M.,
RA   Hernandez-Lucas I., Hong A., Huizar L., Hyman R.W., Jones T., Kahn D.,
RA   Kahn M.L., Kalman S., Keating D.H., Kiss E., Komp C., Lelaure V.,
RA   Masuy D., Palm C., Peck M.C., Pohl T.M., Portetelle D., Purnelle B.,
RA   Ramsperger U., Surzycki R., Thebault P., Vandenbol M.,
RA   Vorhoelter F.J., Weidner S., Wells D.H., Wong K., Yeh K.-C., Batut J.;
RT   "The composite genome of the legume symbiont Sinorhizobium meliloti.";
RL   Science 293:668-672(2001).
DR   EMBL; AL591789; CAC46543.1; -.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   47 AA;  5038 MW;  58EB722379F00CD8 CRC64;


  Query Match             53.8%;  Score 28;  DB 16;  Length 47;
  Best Local Similarity   66.7%;  Pred. No. 4.7e+02;
  Matches    6;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 VVAESPKKP 10
            ||||:|  |
Db      29 VVAEAPVGP 37


RESULT   5
O19452
ID   O19452      PRELIMINARY;      PRT;    25 AA.
AC   O19452;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   MHC CLASS II-ASSOCIATED INVARIANT CHAIN (FRAGMENT).
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90257363; PubMed=2111346;
RA   Eades A.-M., Litfin M., Rahmsdorf H.J.;
RT   "The IFN-gamma response of the murine invariant chain gene is mediated
RT   by a complex enhancer that includes several MHC class II consensus
RT   elements.";
RL   J. Immunol. 144:4399-4409(1990).
DR   EMBL; M35872; AAA37897.1; -.
FT   NON_TER      25     25
SQ   SEQUENCE   25 AA;  2973 MW;  D22EADA1B6036FCC CRC64;


  Query Match             51.9%;  Score 27;  DB 7;  Length 25;
  Best Local Similarity   30.0%;  Pred. No. 3.9e+02;
  Matches    3;  Conservative    4;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 PVVAESPKKP 10
            |::   |::|
Db      15 PILGNRPREP 24


RESULT   6
O88771
ID   O88771      PRELIMINARY;      PRT;    37 AA.
AC   O88771;
```

```
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   26S PROTEASOME P112 SUBUNIT (FRAGMENT).
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=BRAIN;
RX   MEDLINE=98434650; PubMed=9757068;
RA   Seidel B., Keilhoff G., Reinheckel T., Wolf G.;
RT   "Differentially expressed genes in hippocampal cell cultures in
RT   response to an excitotoxic insult by quinolinic acid.";
RL   Mol. Brain Res. 60:296-300(1998).
DR   EMBL; AJ007476; CAA07524.1; -.
KW   Proteasome.
FT   NON_TER       1      1
FT   NON_TER      37     37
SQ   SEQUENCE   37 AA;  4207 MW;  96B6333E4549A626 CRC64;


  Query Match             51.9%;  Score 27;  DB 11;  Length 37;
  Best Local Similarity   62.5%;  Pred. No. 5.7e+02;
  Matches    5;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 PVVAESPK 8
            || |  ||
Db      11 PVAAHGPK 18


RESULT   7
P82570
ID   P82570      PRELIMINARY;      PRT;    39 AA.
AC   P82570;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   HYPOXANTHINE-GUANINE PHOSPHORIBOSYLTRANSFERASE (EC 2.4.2.8) (HGPRT)
DE   (HGPRTASE) (FRAGMENTS).
OS   Streptococcus pyogenes.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC   Streptococcus.
OX   NCBI_TaxID=1314;
RN   [1]
RP   SEQUENCE, AND MASS SPECTROMETRY.
RC   STRAIN=JRS4;
RA   Hogan D.A., Du P., Stevenson T.I., Whitton M., Kilby G.W., Rogers J.,
RA   VanBogelen R.A.;
RT   "Two-dimensional gel electrophoresis map of Streptococcus pyogenes
RT   proteins.";
RL   Submitted (MAY-2000) to the SWISS-PROT data bank.
CC   -!- CATALYTIC ACTIVITY: IMP + PYROPHOSPHATE = HYPOXANTHINE +
CC       5-PHOSPHO-ALPHA-D-RIBOSE 1-DIPHOSPHATE (GUANINE CAN REPLACE
CC       HYPOXANTHINE TO PRODUCE GMP).
CC   -!- PATHWAY: PURINE SALVAGE.
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
CC   -!- MASS SPECTROMETRY: MW=35707.93; METHOD=ELECTROSPRAY.
CC   -!- SIMILARITY: BELONGS TO THE PURINE/PYRIMIDINE
CC       PHOSPHORIBOSYLTRANSFERASE FAMILY.
DR   InterPro; IPR002375; Pur_pyr_pr_transf.
DR   PROSITE; PS00103; PUR_PYR_PR_TRANSFER; FALSE_NEG.
KW   Transferase; Glycosyltransferase; Purine salvage; Magnesium.
FT   NON_TER       1      1
FT   NON_CONS     14     15
FT   NON_CONS     27     28
FT   NON_TER      39     39
SQ   SEQUENCE   39 AA;  4167 MW;  0FE33A096B1BCA1F CRC64;


  Query Match             51.9%;  Score 27;  DB 2;  Length 39;
  Best Local Similarity   75.0%;  Pred. No. 6e+02;
  Matches    6;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 VVAESPKK 9
            ||  ||||
Db       8 VVVVSPKK 15


RESULT   8
Q95R01
ID   Q95R01      PRELIMINARY;      PRT;    42 AA.
AC   Q95R01;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HYPOTHETICAL 4.7 KDA PROTEIN.
GN   K07B1.6.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RX   MEDLINE=99069613; PubMed=9851916;
RA   None;
RT   "Genome sequence of the nematode C. elegans: a platform for
RT   investigating biology. The C. elegans Sequencing Consortium.";
RL   Science 282:2012-2018(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   Pauley A., Gattung S.;
RT   "The sequence of C. elegans cosmid K07B1.";
RL   Submitted (MAY-1997) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   Waterston R.;
RT   "Direct Submission.";
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF003384; AAL08041.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   42 AA;  4709 MW;  390F7ADF3BF906E9 CRC64;


  Query Match             51.9%;  Score 27;  DB 5;  Length 42;
  Best Local Similarity   66.7%;  Pred. No. 6.4e+02;
  Matches    4;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5 ESPKKP 10
            :||:||
Db      11 QSPRKP 16


RESULT   9
Q9JIC8
ID   Q9JIC8      PRELIMINARY;      PRT;    47 AA.
AC   Q9JIC8;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   SMN PROTEIN (FRAGMENT).
GN   SMN.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J;
RX   MEDLINE=20414754; PubMed=10958634;
RA   Growney J.D., Dietrich W.F.;
RT   "High-resolution genetic and physical map of the Lgn1 interval in
```

```
RT   C57BL/6J implicates Naip2 or Naip5 in Legionella pneumophila
RT   pathogenesis.";
RL   Genome Res. 10:1158-1171(2000).
DR   EMBL; AF240503; AAF81197.1; -.
DR   MGD; MGI:109257; Smn.
FT   NON_TER       1      1
FT   NON_TER      47     47
SQ   SEQUENCE   47 AA;  5160 MW;  1CD41D6E32BC7126 CRC64;


  Query Match             51.9%;  Score 27;  DB 11;  Length 47;
  Best Local Similarity   50.0%;  Pred. No. 7.2e+02;
  Matches    4;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      3 VAESPKKP 10
          : |:| ||
Db     15 ICETPDKP 22


RESULT  10
Q9GB18
ID   Q9GB18      PRELIMINARY;      PRT;    29 AA.
AC   Q9GB18;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   CYTOCHROME C OXIDASE SUBUNIT II (FRAGMENT).
GN   COII.
OS   Melidectes belfordi (Belford's melidectes).
OG   Mitochondrion.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Passeriformes; Meliphagidae;
OC   Melidectes.
OX   NCBI_TaxID=43161;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Slikas B., Jones I.B., Derrickson S.R., Fleischer R.C.;
RT   "Phylogenetic relationships of Micronesian white-eyes (Zosteropidae)
RT   based on mitochondrial sequence data.";
RL   Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF168458; AAG12314.1; -.
DR   InterPro; IPR002429; Cyt_c_ox_2.
DR   Pfam; PF00116; COX2; 1.
KW   Mitochondrion.
FT   NON_TER       1      1
SQ   SEQUENCE   29 AA;  3198 MW;  43F35FDA214CBD79 CRC64;


  Query Match             50.0%;  Score 26;  DB 8;  Length 29;
  Best Local Similarity   50.0%;  Pred. No. 6.8e+02;
  Matches    5;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 PVVAESPKKP 10
          |:| ||   |
Db      7 PIVVESAPLP 16


RESULT  11
Q84028
ID   Q84028      PRELIMINARY;      PRT;    47 AA.
AC   Q84028;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   INFLUENZA A/PUERTO RICO/8/34 (H0N1), MATRIX PROTEIN (SEG 7), RNA, 5'
DE   END (FRAGMENT).
OS   Influenza A virus.
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses.
OX   NCBI_TaxID=11320;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=79213454; PubMed=572297;
RA   Both G.W., Air G.M.;
RT   "Nucleotide sequence coding for the N-terminal region of the matrix
RT   protein of influenza virus.";
RL   Eur. J. Biochem. 96:363-372(1979).
DR   EMBL; M10642; AAA43309.1; -.
DR   HSSP; P03485; 1AA7.
DR   InterPro; IPR001561; Flu_M1.
DR   Pfam; PF00598; Flu_M1; 1.
DR   ProDom; PD001061; Flu_M1; 1.
KW   Matrix protein.
FT   NON_TER      47     47
SQ   SEQUENCE   47 AA;  5343 MW;  83415B57C9CE4D51 CRC64;


  Query Match             50.0%;  Score 26;  DB 12;  Length 47;
  Best Local Similarity   44.4%;  Pred. No. 1.1e+03;
  Matches    4;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 VVAESPKKP 10
          :|  :| ||
Db     14 IVPSAPSKP 22


RESULT  12
Q96L04
ID   Q96L04      PRELIMINARY;      PRT;    18 AA.
AC   Q96L04;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HYPOTHETICAL 2.1 KDA PROTEIN (FRAGMENT).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=TESTIS;
RA   Strausberg R.;
RL   Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC014606; AAH14606.1; -.
KW   Hypothetical protein.
FT   NON_TER       1      1
SQ   SEQUENCE   18 AA;  2097 MW;  5165856AB840D288 CRC64;


  Query Match             48.1%;  Score 25;  DB 4;  Length 18;
  Best Local Similarity   57.1%;  Pred. No. 6.5e+02;
  Matches    4;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      4 AESPKKP 10
          || |: |
Db     12 AEGPRSP 18


RESULT  13
Q16070
ID   Q16070      PRELIMINARY;      PRT;    20 AA.
AC   Q16070;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   NEUROFILAMENT HEAVY SUBUNIT (FRAGMENT).
GN   NEFH.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94040777; PubMed=8224877;
RA   Figlewicz D.A., Rouleau G.A., Krizus A., Julien J.P.;
RT   "Polymorphism in the multi-phosphorylation domain of the human
```

```
RT   neurofilament heavy-subunit-encoding gene.";
RL   Gene 132:297-300(1993).
DR   EMBL; S66488; AAB28609.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   20 AA;  2198 MW;  E9A0975B41FD8082 CRC64;


  Query Match             48.1%;  Score 25;  DB 4;  Length 20;
  Best Local Similarity   55.6%;  Pred. No. 7.2e+02;
  Matches    5;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 PVVAESPKK 9
          |  |:|||:|
Db      2 PEKAKSPEK 10


RESULT  14
Q9SLV6
ID   Q9SLV6      PRELIMINARY;      PRT;    22 AA.
AC   Q9SLV6;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   NITRILASE-LIKE PROTEIN (FRAGMENT).
GN   TNIT4B.
OS   Nicotiana tabacum (Common tobacco).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
OX   NCBI_TaxID=4097;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SR-1;
RX   MEDLINE=20039615; PubMed=10574458;
RA   Dohmoto M., Sano J., Tsunoda H., Yamaguchi K.;
RT   "Structural analysis of the TNIT4 genes encoding nitrilase-like
RT   protein from tobacco.";
RL   DNA Res. 6:313-317(1999).
DR   EMBL; AB027127; BAA77683.1; -.
FT   NON_TER      22     22
SQ   SEQUENCE   22 AA;  2320 MW;  ACB9F59BC2461022 CRC64;


  Query Match             48.1%;  Score 25;  DB 10;  Length 22;
  Best Local Similarity   71.4%;  Pred. No. 7.9e+02;
  Matches    5;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 PVVAESP 7
          ||| | |
Db      7 PVVNEGP 13


RESULT  15
Q9UMH7
ID   Q9UMH7      PRELIMINARY;      PRT;    27 AA.
AC   Q9UMH7;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   SS-B/LA PROTEIN (FRAGMENT).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90237237; PubMed=1692037;
RA   Kohsaka H., Yamamoto K., Fujii H., Miura H., Miyasaka N., Nishioka K.,
RA   Miyamoto T.;
RT   "Fine epitope mapping the human SS-B/La protein: Identification of a
RT   distinct autoepitope homologous to a viral gag polyprotein.";
RL   J. Clin. Invest. 85:1566-1574(1990).
DR   EMBL; M35261; AAA36652.1; -.
FT   NON_TER       1      1
FT   NON_TER      27     27
SQ   SEQUENCE   27 AA;  3175 MW;  A9B8FFD9F097F035 CRC64;


  Query Match             48.1%;  Score 25;  DB 4;  Length 27;
  Best Local Similarity   50.0%;  Pred. No. 9.7e+02;
  Matches    4;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      3 VAESPKKP 10
          :  || ||
Db      9 IRRSPSKP 16


Search completed: October 12, 2002, 20:50:23
Job time: 187 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: October 12, 2002, 16:51:25 ; Search time 60.75 Seconds
(without alignments)
18.284 Million cell updates/sec

Title: US-09-408-578A-1
Perfect score: 52
Sequence: 1 PVVAESPKKP 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802:*
1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	52	100.0	10	18	AAW41844	Modified B. burgdo
2	52	100.0	10	18	AAW41821	B. burgdorferi sen
3	52	100.0	11	18	AAW41825	Modified B. burgdo
4	52	100.0	15	18	AAW41827	B. burgdorferi sen
5	52	100.0	20	18	AAW41826	B. burgdorferi sen
6	52	100.0	23	20	AAY27428	Borrelia outer sur
7	52	100.0	24	20	AAY27429	Borrelia outer sur
8	52	100.0	191	15	AAR60897	Borrelia VS461 ant
9	52	100.0	192	15	AAR60889	Borrelia 297 antig
10	52	100.0	192	15	AAR60886	Borrelia IP2 OspC
11	52	100.0	193	15	AAR60893	Borrelia ACA1 anti
12	52	100.0	194	15	AAR60894	Borrelia H9 antige
13	52	100.0	194	15	AAR60896	Borrelia JSB antig
14	52	100.0	194	21	AAY78908	Outer surface prot
15	52	100.0	206	22	AAB62723	B afzelii ospC pro
16	52	100.0	207	16	AAR75730	B. burgdorferi str
17	52	100.0	207	18	AAW41823	B. garinii OspC.
18	52	100.0	209	16	AAR75728	B. burgdorferi str
19	52	100.0	209	22	AAB62720	B burgdorferi ospC
20	52	100.0	210	16	AAW11935	Outer surface prot
21	52	100.0	210	16	AAR75727	B. burgdorferi str
22	52	100.0	211	22	AAB62722	B garinii ospC pro
23	52	100.0	212	12	AAR13140	B.burgdorferi stra
24	52	100.0	212	16	AAR75729	B. burgdorferi str
25	52	100.0	212	18	AAW41824	B. afzelii OspC.
26	52	100.0	377	22	AAB62713	Borrelia sp chimer
27	52	100.0	378	22	AAB62712	Borrelia sp chimer
28	52	100.0	378	22	AAB62725	Borrelia sp chimer
29	52	100.0	384	22	AAB62726	Borrelia sp chimer
30	52	100.0	386	22	AAB62727	Borrelia sp chimer
31	52	100.0	400	22	AAB62739	Borrelia sp chimer
32	52	100.0	401	22	AAB62733	Borrelia sp chimer
33	52	100.0	401	22	AAB62738	Borrelia sp chimer
34	52	100.0	408	22	AAB62737	Borrelia sp chimer
35	52	100.0	410	22	AAB62740	Borrelia sp chimer
36	52	100.0	466	16	AAR75739	B31 outer surface
37	52	100.0	466	16	AAR75740	B31 outer surface
38	52	100.0	587	16	AAR75746	B31 Osp-A/antigen
39	51	98.1	191	15	AAR60884	Borrelia 28691 Osp
40	49	94.2	10	18	AAW41838	Modified B. burgdo
41	49	94.2	11	15	AAR70367	Borrelia OspC anti
42	49	94.2	15	15	AAR70362	Borrelia OspC anti
43	49	94.2	189	15	AAR60907	Borrelia KL11 anti
44	49	94.2	189	15	AAR60909	Borrelia H13 antig
45	49	94.2	191	15	AAR60898	Borrelia M57 antig

ALIGNMENTS

RESULT 1
AAW41844
ID AAW41844 standard; peptide; 10 AA.
XX
AC AAW41844;
XX
DT 14-MAY-1998 (first entry)
XX
DE Modified B. burgdorferi sensu lato OspC C-terminal peptide.
XX
KW Sensu lato; outer surface protein C; OspC; diagnosis; Lyme disease;
KW vaccine; infection.
XX
OS Borrelia burgdorferi.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 10
FT /label= amidated
XX
PN WO9742221-A1.
XX
PD 13-NOV-1997.
XX
PF 02-MAY-1997; 97WO-DK00203.
XX
PR 02-MAY-1996; 96DK-0000526.
XX
PA (STAT-) STATENS SERUMINSTITUT.
XX
PI Holm A, Mathiesen MJ, Ostergaard S, Theisen M;
XX
DR WPI; 1997-558908/51.
XX

PT Detecting previous sensitisation to the OspC protein of Borellia
PT burgdorferi - by detecting immunoreactivity between patient T cells
PT or immunoglobulins and C-terminal peptide of the protein
XX
PS Example 3; Page 53; 95pp; English.
XX
CC The present sequence was used in the development of a novel method
CC for the identification of a patient's previous sensitisation to
CC Borellia burgdorferi sensu lato outer surface protein C (OspC).
CC The method comprises reacting immunoglobulin (Ig) or T cells from
CC the patient with a polypeptide of at most 60 amino acids containing
CC a peptide with at least 50% identity to the B. burgdorferi derived
CC sequence AAW41821, or its subsequences of at least 5 amino acids. The
CC degree of immunological reactivity between the polypeptide and Ig
CC or T cells is measured and significant reactivity is indicative of
CC sensitisation.
CC The method can be used to diagnose Lyme disease and is based on
CC reactivity with antibodies against the OspC protein. The test can
CC be done in vitro or in vivo, e.g. as a skin test. Vaccine
CC compositions comprising the polypeptide can be used to protect
CC humans and other animals against B. burgdorferi infection. The
CC polypeptide provides higher sensitivity than full-length OspC, and
CC so is better at detecting infection in its early stages, especially
CC when combined with the known assay for flagellar proteins. The
CC seven carboxy-terminal residues of AAW41821 represent an epitope
CC essential for human immune response to OspC. The polypeptide is
CC also easier to prepare and purify than (nearly) full-length
CC protein, facilitating standardisation of the assay, and is less
CC cross-reactive with antibodies raised against other antigens. The
CC small size of the polypeptide allows a high density of binding
CC sites to be created on a solid support. Incorporation of
CC non-natural amino acid into the polypeptide increases its
CC resistance to peptidases when used in vivo.
XX
SQ Sequence 10 AA;

Query Match 100.0%; Score 52; DB 18; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0078;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PVVAESPKKP 10
||||||||||
Db 1 pvvaespkkp 10

RESULT 2
AAW41821
ID AAW41821 standard; peptide; 10 AA.
XX
AC AAW41821;
XX
DT 14-MAY-1998 (first entry)
XX
DE B. burgdorferi sensu lato OspC carboxy-terminal peptide.
XX
KW Sensu lato; outer surface protein C; OspC; diagnosis; Lyme disease;
KW vaccine; infection.
XX
OS Borrelia burgdorferi.
XX
PN WO9742221-A1.
XX
PD 13-NOV-1997.
XX
PF 02-MAY-1997; 97WO-DK00203.
XX
PR 02-MAY-1996; 96DK-0000526.
XX
PA (STAT-) STATENS SERUMINSTITUT.
XX
PI Holm A, Mathiesen MJ, Ostergaard S, Theisen M;
XX
DR WPI; 1997-558908/51.
XX
PT Detecting previous sensitisation to the OspC protein of Borellia
PT burgdorferi - by detecting immunoreactivity between patient T cells
PT or immunoglobulins and C-terminal peptide of the protein
XX
PS Claim 1; Page 77; 95pp; English.
XX
CC The present sequence was used in the development of a novel method
CC for the identification of a patient's previous sensitisation to
CC Borellia burgdorferi sensu lato outer surface protein C (OspC).
CC The method comprises reacting immunoglobulin (Ig) or T cells from
CC the patient with a polypeptide of at most 60 amino acids containing
CC a peptide with at least 50% identity to the B. burgdorferi derived
CC sequence AAW41821, or its subsequences of at least 5 amino acids. The
CC degree of immunological reactivity between the polypeptide and Ig
CC or T cells is measured and significant reactivity is indicative of
CC sensitisation.
CC The method can be used to diagnose Lyme disease and is based on
CC reactivity with antibodies against the OspC protein. The test can
CC be done in vitro or in vivo, e.g. as a skin test. Vaccine
CC compositions comprising the polypeptide can be used to protect
CC humans and other animals against B. burgdorferi infection. The
CC polypeptide provides higher sensitivity than full-length OspC, and
CC so is better at detecting infection in its early stages, especially
CC when combined with the known assay for flagellar proteins. The
CC seven carboxy-terminal residues of AAW41821 represent an epitope
CC essential for human immune response to OspC. The polypeptide is
CC also easier to prepare and purify than (nearly) full-length
CC protein, facilitating standardisation of the assay, and is less
CC cross-reactive with antibodies raised against other antigens. The
CC small size of the polypeptide allows a high density of binding
CC sites to be created on a solid support. Incorporation of
CC non-natural amino acid into the polypeptide increases its
CC resistance to peptidases when used in vivo.
XX
SQ Sequence 10 AA;

Query Match 100.0%; Score 52; DB 18; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0078;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PVVAESPKKP 10
||||||||||
Db 1 pvvaespkkp 10

RESULT 3
AAW41825
ID AAW41825 standard; peptide; 11 AA.
XX
AC AAW41825;
XX
DT 14-MAY-1998 (first entry)
XX
DE Modified B. burgdorferi sensu lato OspC C-terminal peptide.
XX
KW Sensu lato; outer surface protein C; OspC; diagnosis; Lyme disease;
KW vaccine; infection.
XX
OS Borrelia burgdorferi.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /note= "6-aminohexanoic acid"
XX
PN WO9742221-A1.
XX
PD 13-NOV-1997.
XX
PF 02-MAY-1997; 97WO-DK00203.

XX
PR 02-MAY-1996; 96DK-0000526.
XX
PA (STAT-) STATENS SERUMINSTITUT.
XX
PI Holm A, Mathiesen MJ, Ostergaard S, Theisen M;
XX
DR WPI; 1997-558908/51.
XX
PT Detecting previous sensitisation to the OspC protein of Borellia
PT burgdorferi - by detecting immunoreactivity between patient T cells
PT or immunoglobulins and C-terminal peptide of the protein
XX
PS Example; Page 40; 95pp; English.
XX
CC The present sequence was used in the development of a novel method
CC for the identification of a patient's previous sensitisation to
CC Borellia burgdorferi sensu lato outer surface protein C (OspC).
CC The method comprises reacting immunoglobulin (Ig) or T cells from
CC the patient with a polypeptide of at most 60 amino acids containing
CC a peptide with at least 50% identity to the B. burgdorferi derived
CC sequence AAW41821, or its subsequences of at least 5 amino acids. The
CC degree of immunological reactivity between the polypeptide and Ig
CC or T cells is measured and significant reactivity is indicative of
CC sensitisation.
CC The method can be used to diagnose Lyme disease and is based on
CC reactivity with antibodies against the OspC protein. The test can
CC be done in vitro or in vivo, e.g. as a skin test. Vaccine
CC compositions comprising the polypeptide can be used to protect
CC humans and other animals against B. burgdorferi infection. The
CC polypeptide provides higher sensitivity than full-length OspC, and
CC so is better at detecting infection in its early stages, especially
CC when combined with the known assay for flagellar proteins. The
CC seven carboxy-terminal residues of AAW41821 represent an epitope
CC essential for human immune response to OspC. The polypeptide is
CC also easier to prepare and purify than (nearly) full-length
CC protein, facilitating standardisation of the assay, and is less
CC cross-reactive with antibodies raised against other antigens. The
CC small size of the polypeptide allows a high density of binding
CC sites to be created on a solid support. Incorporation of
CC non-natural amino acid into the polypeptide increases its
CC resistance to peptidases when used in vivo.
XX
SQ Sequence 11 AA;

Query Match 100.0%; Score 52; DB 18; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.0085;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PVVAESPKKP 10
||||||||||
Db 2 pvvaespkkp 11

RESULT 4
AAW41827
ID AAW41827 standard; peptide; 15 AA.
XX
AC AAW41827;
XX
DT 14-MAY-1998 (first entry)
XX
DE B. burgdorferi sensu lato OspC carboxy-terminal peptide.
XX
KW Sensu lato; outer surface protein C; OspC; diagnosis; Lyme disease;
KW vaccine; infection.
XX
OS Borrelia burgdorferi.
XX
PN WO9742221-A1.
XX
PD 13-NOV-1997.
XX
PF 02-MAY-1997; 97WO-DK00203.
XX
PR 02-MAY-1996; 96DK-0000526.
XX
PA (STAT-) STATENS SERUMINSTITUT.
XX
PI Holm A, Mathiesen MJ, Ostergaard S, Theisen M;
XX
DR WPI; 1997-558908/51.
XX
PT Detecting previous sensitisation to the OspC protein of Borellia
PT burgdorferi - by detecting immunoreactivity between patient T cells
PT or immunoglobulins and C-terminal peptide of the protein
XX
PS Example 3; Page 51; 95pp; English.
XX
CC The present sequence was used in the development of a novel method
CC for the identification of a patient's previous sensitisation to
CC Borellia burgdorferi sensu lato outer surface protein C (OspC).
CC The method comprises reacting immunoglobulin (Ig) or T cells from
CC the patient with a polypeptide of at most 60 amino acids containing
CC a peptide with at least 50% identity to the B. burgdorferi derived
CC sequence AAW41821, or its subsequences of at least 5 amino acids. The
CC degree of immunological reactivity between the polypeptide and Ig
CC or T cells is measured and significant reactivity is indicative of
CC sensitisation.
CC The method can be used to diagnose Lyme disease and is based on
CC reactivity with antibodies against the OspC protein. The test can
CC be done in vitro or in vivo, e.g. as a skin test. Vaccine
CC compositions comprising the polypeptide can be used to protect
CC humans and other animals against B. burgdorferi infection. The
CC polypeptide provides higher sensitivity than full-length OspC, and
CC so is better at detecting infection in its early stages, especially
CC when combined with the known assay for flagellar proteins. The
CC seven carboxy-terminal residues of AAW41821 represent an epitope
CC essential for human immune response to OspC. The polypeptide is
CC also easier to prepare and purify than (nearly) full-length
CC protein, facilitating standardisation of the assay, and is less
CC cross-reactive with antibodies raised against other antigens. The
CC small size of the polypeptide allows a high density of binding
CC sites to be created on a solid support. Incorporation of
CC non-natural amino acid into the polypeptide increases its
CC resistance to peptidases when used in vivo.
XX
SQ Sequence 15 AA;

Query Match 100.0%; Score 52; DB 18; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.012;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PVVAESPKKP 10
||||||||||
Db 6 pvvaespkkp 15

RESULT 5
AAW41826
ID AAW41826 standard; peptide; 20 AA.
XX
AC AAW41826;
XX
DT 14-MAY-1998 (first entry)
XX
DE B. burgdorferi sensu lato OspC carboxy-terminal peptide.
XX
KW Sensu lato; outer surface protein C; OspC; diagnosis; Lyme disease;
KW vaccine; infection.
XX
OS Borrelia burgdorferi.
XX
PN WO9742221-A1.

XX
PD 13-NOV-1997.
XX
PF 02-MAY-1997; 97WO-DK00203.
XX
PR 02-MAY-1996; 96DK-0000526.
XX
PA (STAT-) STATENS SERUMINSTITUT.
XX
PI Holm A, Mathiesen MJ, Ostergaard S, Theisen M;
XX
DR WPI; 1997-558908/51.
XX
PT Detecting previous sensitisation to the OspC protein of Borellia
PT burgdorferi - by detecting immunoreactivity between patient T cells
PT or immunoglobulins and C-terminal peptide of the protein
XX
PS Example 3; Page 51; 95pp; English.
XX
CC The present sequence was used in the development of a novel method
CC for the identification of a patient's previous sensitisation to
CC Borellia burgdorferi sensu lato outer surface protein C (OspC).
CC The method comprises reacting immunoglobulin (Ig) or T cells from
CC the patient with a polypeptide of at most 60 amino acids containing
CC a peptide with at least 50% identity to the B. burgdorferi derived
CC sequence AAW41821, or its subsequences of at least 5 amino acids. The
CC degree of immunological reactivity between the polypeptide and Ig
CC or T cells is measured and significant reactivity is indicative of
CC sensitisation.
CC The method can be used to diagnose Lyme disease and is based on
CC reactivity with antibodies against the OspC protein. The test can
CC be done in vitro or in vivo, e.g. as a skin test. Vaccine
CC compositions comprising the polypeptide can be used to protect
CC humans and other animals against B. burgdorferi infection. The
CC polypeptide provides higher sensitivity than full-length OspC, and
CC so is better at detecting infection in its early stages, especially
CC when combined with the known assay for flagellar proteins. The
CC seven carboxy-terminal residues of AAW41821 represent an epitope
CC essential for human immune response to OspC. The polypeptide is
CC also easier to prepare and purify than (nearly) full-length
CC protein, facilitating standardisation of the assay, and is less
CC cross-reactive with antibodies raised against other antigens. The
CC small size of the polypeptide allows a high density of binding
CC sites to be created on a solid support. Incorporation of
CC non-natural amino acid into the polypeptide increases its
CC resistance to peptidases when used in vivo.
XX
SQ Sequence 20 AA;

Query Match 100.0%; Score 52; DB 18; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.016;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PVVAESPKKP 10
||||||||||
Db 11 pvvaespkkp 20

RESULT 6
AAY27428
ID AAY27428 standard; peptide; 23 AA.
XX
AC AAY27428;
XX
DT 26-NOV-1999 (first entry)
XX
DE Borrelia outer surface protein C (OspC) C-terminal peptide fragment.
XX
KW Borrelia; IgM antibody; outer surface protein C; OspC; deer tick;
KW cerebrospinal fluid; Lyme borreliosis; micro-capture assay; flagellum;
KW epitope.
XX
OS Synthetic.
OS Borrelia burgdorferi.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /note= "linked to biotin via an O-linker of formula
FT [2-aminoethoxy]ethoxy acetic acid"
XX
PN EP949508-A1.
XX
PD 13-OCT-1999.
XX
PF 07-APR-1999; 99EP-0610026.
XX
PR 08-APR-1998; 98DK-0000516.
XX
PA (DAKO-) DAKO AS.
XX
PI Staffeldt Schou O, Winther L, Stender H;
XX
DR WPI; 1999-553537/47.
XX
PT Diagnosing Lyme borreliosis by detecting antibodies against two
PT antigens simultaneously -
XX
PS Example 1; Page 7; 23pp; English.
XX
CC The invention provides a new method for detecting IgM antibodies against
CC Borrelia burgdorferi in a sample of human or animal fluid. The method
CC comprises: (1) contacting antibodies in the sample with anti-IgM
CC immobilized to a solid support, (2) separating the support from the
CC liquid phase; and (3) contacting the bound antibodies with a complex
CC comprising at least one set of B. burgdorferi (outer surface protein C
CC (OspC) peptides and/or at least one set of other B. burgdorferi peptides,
CC each attached to a carrier; and (4) detecting the presence of antibodies
CC against B. burgdorferi. The new method may be used to detect antibodies
CC against B. burgdorferi in (especially) serum or cerebrospinal fluid
CC samples from patients bitten by deer ticks. B. burgdorferi causes Lyme
CC borreliosis so detection of antibodies against it allows diagnosis of
CC infection by this organism. The method is a micro-capture assay in which
CC the antigen complex is a combination of the B.burgdorferi flagellum and
CC OspC peptides. The presence of epitopes from both antigens in the complex
CC allows the simultaneous detection of serum antibodies against these
CC proteins which increases the sensitivity of the test. The two antigens
CC are pure, which also decreases the possibility of cross-reactivity. The
CC present sequence represents a Borrelia OspC C-terminal peptide fragment.
XX
SQ Sequence 23 AA;

Query Match 100.0%; Score 52; DB 20; Length 23;
Best Local Similarity 100.0%; Pred. No. 0.018;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PVVAESPKKP 10
||||||||||
Db 14 pvvaespkkp 23

RESULT 7
AAY27429
ID AAY27429 standard; peptide; 24 AA.
XX
AC AAY27429;
XX
DT 26-NOV-1999 (first entry)
XX
DE Borrelia outer surface protein C (OspC) C-terminal peptide fragment.
XX
KW Borrelia; IgM antibody; outer surface protein C; OspC; deer tick;
KW cerebrospinal fluid; Lyme borreliosis; micro-capture assay; flagellum;
KW epitope.
XX

OS Synthetic.
OS Borrelia burgdorferi.
XX
PN EP949508-A1.
XX
PD 13-OCT-1999.
XX
PF 07-APR-1999; 99EP-0610026.
XX
PR 08-APR-1998; 98DK-0000516.
XX
PA (DAKO-) DAKO AS.
XX
PI Staffeldt Schou O, Winther L, Stender H;
XX
DR WPI; 1999-553537/47.
XX
PT Diagnosing Lyme borreliosis by detecting antibodies against two
PT antigens simultaneously -
XX
PS Example 1; Page 7; 23pp; English.
XX
CC The invention provides a new method for detecting IgM antibodies against
CC Borrelia burgdorferi in a sample of human or animal fluid. The method
CC comprises: (1) contacting antibodies in the sample with anti-IgM
CC immobilized to a solid support, (2) separating the support from the
CC liquid phase; and (3) contacting the bound antibodies with a complex
CC comprising at least one set of B. burgdorferi (outer surface protein C
CC (OspC) peptides and/or at least one set of other B. burgdorferi peptides,
CC each attached to a carrier; and (4) detecting the presence of antibodies
CC against B. burgdorferi. The new method may be used to detect antibodies
CC against B. burgdorferi in (especially) serum or cerebrospinal fluid
CC samples from patients bitten by deer ticks. B. burgdorferi causes Lyme
CC borreliosis so detection of antibodies against it allows diagnosis of
CC infection by this organism. The method is a micro-capture assay in which
CC the antigen complex is a combination of the B.burgdorferi flagellum and
CC OspC peptides. The presence of epitopes from both antigens in the complex
CC allows the simultaneous detection of serum antibodies against these
CC proteins which increases the sensitivity of the test. The two antigens
CC are pure, which also decreases the possibility of cross-reactivity. The
CC present sequence represents a Borrelia OspC C-terminal peptide fragment,
CC where the N-terminal cysteine residue has been incorporated to provide a
CC SH group to be used in a coupling reaction.
XX
SQ Sequence 24 AA;

Query Match 100.0%; Score 52; DB 20; Length 24;
Best Local Similarity 100.0%; Pred. No. 0.019;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PVVAESPKKP 10
||||||||||
Db 15 pvvaespkkp 24

RESULT 8
AAR60897
ID AAR60897 standard; Protein; 191 AA.
XX
AC AAR60897;
XX
DT 25-MAY-1995 (first entry)
XX
DE Borrelia VS461 antigen vaccine.
XX
KW OspC antigen; vaccine; Lyme disease; borreliosis; immunogen;
KW serovar typing; restriction fragment length polymorphism;
KW RFLP; Pichia pastoris.
XX
OS Borrelia burgdorferi VS461.
XX
PN WO9425596-A.
XX
PD 10-NOV-1994.
XX
PF 29-APR-1994; 94WO-EP01365.
XX
PR 29-APR-1993; 93US-0053863.
XX
PA (IMMO) IMMUNO AG.
XX
PI Crowe B, Dorner F, Livey I;
XX
DR WPI; 1994-358273/44.
DR N-PSDB; AAQ73870.
XX
PT Immunogenic composition comprising OspC antigens - for the
PT treatment of Lyme borreliosis in different, specific geographical
PT areas.
XX
PS Disclosure; Fig. 9a; 115pp; English.
XX
CC A vaccine for Lyme disease includes selected OspC antigen
CC formulations based on defined OspC families resolved by serovar
CC typing and RFLP typing. Partial sequences of OspC genes selected
CC from different RFLP types are given in AAQ73883-905 (encoded peptides,
CC comprising the first 92% of mature OspC, are given in AAR62771-93).
CC Complete sequences of these novel ospC genes, including the 3' end,
CC plus sequences for the ospC genes of Borrelia strains H13 and 28691
CC are given in AAQ73857-82, and encoded proteins in AAR60884-909. The
CC DNA sequences may be expressed in e.g. Pichia pastoris for
CC recombinant antigen production.
XX
SQ Sequence 191 AA;

Query Match 100.0%; Score 52; DB 15; Length 191;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PVVAESPKKP 10
||||||||||
Db 182 pvvaespkkp 191

RESULT 9
AAR60889
ID AAR60889 standard; Protein; 192 AA.
XX
AC AAR60889;
XX
DT 25-MAY-1995 (first entry)
XX
DE Borrelia 297 antigen vaccine.
XX
KW OspC antigen; vaccine; Lyme disease; borreliosis; immunogen;
KW serovar typing; restriction fragment length polymorphism;
KW RFLP; Pichia pastoris.
XX
OS Borrelia burgdorferi 297.
XX
PN WO9425596-A.
XX
PD 10-NOV-1994.
XX
PF 29-APR-1994; 94WO-EP01365.
XX
PR 29-APR-1993; 93US-0053863.
XX
PA (IMMO) IMMUNO AG.
XX
PI Crowe B, Dorner F, Livey I;
XX
DR WPI; 1994-358273/44.
DR N-PSDB; AAQ73862.

XX
PT Immunogenic composition comprising OspC antigens - for the
PT treatment of Lyme borreliosis in different, specific geographical
PT areas.
XX
PS Disclosure; Fig. 9a; 115pp; English.
XX
CC A vaccine for Lyme disease includes selected OspC antigen
CC formulations based on defined OspC families resolved by serovar
CC typing and RFLP typing. Partial sequences of OspC genes selected
CC from different RFLP types are given in AAQ73883-905 (encoded peptides,
CC comprising the first 92% of mature OspC, are given in AAR62771-93).
CC Complete sequences of these novel ospC genes, including the 3' end,
CC plus sequences for the ospC genes of Borrelia strains H13 and 28691
CC are given in AAQ73857-82, and encoded proteins in AAR60884-909. The
CC DNA sequences may be expressed in e.g. Pichia pastoris for
CC recombinant antigen production.
XX
SQ Sequence 192 AA;

Query Match 100.0%; Score 52; DB 15; Length 192;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PVVAESPKKP 10
||||||||||
Db 183 pvvaespkkp 192

RESULT 10
AAR60886
ID AAR60886 standard; Protein; 192 AA.
XX
AC AAR60886;
XX
DT 25-MAY-1995 (first entry)
XX
DE Borrelia IP2 OspC antigen vaccine.
XX
KW OspC antigen; vaccine; Lyme disease; borreliosis; immunogen;
KW serovar typing; restriction fragment length polymorphism;
KW RFLP; Pichia pastoris.
XX
OS Borrelia burgdorferi IP2.
XX
PN WO9425596-A.
XX
PD 10-NOV-1994.
XX
PF 29-APR-1994; 94WO-EP01365.
XX
PR 29-APR-1993; 93US-0053863.
XX
PA (IMMO) IMMUNO AG.
XX
PI Crowe B, Dorner F, Livey I;
XX
DR WPI; 1994-358273/44.
DR N-PSDB; AAQ73859.
XX
PT Immunogenic composition comprising OspC antigens - for the
PT treatment of Lyme borreliosis in different, specific geographical
PT areas.
XX
PS Disclosure; Fig. 9a; 115pp; English.
XX
CC A vaccine for Lyme disease includes selected OspC antigen
CC formulations based on defined OspC families resolved by serovar
CC typing and RFLP typing. Partial sequences of OspC genes selected
CC from different RFLP types are given in AAQ73883-905 (encoded peptides,
CC comprising the first 92% of mature OspC, are given in AAR62771-93).
CC Complete sequences of these novel ospC genes, including the 3' end,
CC plus sequences for the ospC genes of Borrelia strains H13 and 28691
CC are given in AAQ73857-82, and encoded proteins in AAR60884-909. The
CC DNA sequences may be expressed in e.g. Pichia pastoris for
CC recombinant antigen production.
XX
SQ Sequence 192 AA;

Query Match 100.0%; Score 52; DB 15; Length 192;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PVVAESPKKP 10
||||||||||
Db 183 pvvaespkkp 192

RESULT 11
AAR60893
ID AAR60893 standard; Protein; 193 AA.
XX
AC AAR60893;
XX
DT 25-MAY-1995 (first entry)
XX
DE Borrelia ACA1 antigen vaccine.
XX
KW OspC antigen; vaccine; Lyme disease; borreliosis; immunogen;
KW serovar typing; restriction fragment length polymorphism;
KW RFLP; Pichia pastoris.
XX
OS Borrelia burgdorferi ACA1.
XX
PN WO9425596-A.
XX
PD 10-NOV-1994.
XX
PF 29-APR-1994; 94WO-EP01365.
XX
PR 29-APR-1993; 93US-0053863.
XX
PA (IMMO) IMMUNO AG.
XX
PI Crowe B, Dorner F, Livey I;
XX
DR WPI; 1994-358273/44.
DR N-PSDB; AAQ73866.
XX
PT Immunogenic composition comprising OspC antigens - for the
PT treatment of Lyme borreliosis in different, specific geographical
PT areas.
XX
PS Disclosure; Fig. 9a; 115pp; English.
XX
CC A vaccine for Lyme disease includes selected OspC antigen
CC formulations based on defined OspC families resolved by serovar
CC typing and RFLP typing. Partial sequences of OspC genes selected
CC from different RFLP types are given in AAQ73883-905 (encoded peptides,
CC comprising the first 92% of mature OspC, are given in AAR62771-93).
CC Complete sequences of these novel ospC genes, including the 3' end,
CC plus sequences for the ospC genes of Borrelia strains H13 and 28691
CC are given in AAQ73857-82, and encoded proteins in AAR60884-909. The
CC DNA sequences may be expressed in e.g. Pichia pastoris for
CC recombinant antigen production.
XX
SQ Sequence 193 AA;

Query Match 100.0%; Score 52; DB 15; Length 193;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PVVAESPKKP 10

```
          ||||||||||
Db    184 pvvaespkkp 193


RESULT  12
AAR60894
ID   AAR60894 standard; Protein; 194 AA.
XX
AC   AAR60894;
XX
DT   25-MAY-1995  (first entry)
XX
DE   Borrelia H9 antigen vaccine.
XX
KW   OspC antigen; vaccine; Lyme disease; borreliosis; immunogen;
KW   serovar typing; restriction fragment length polymorphism;
KW   RFLP; Pichia pastoris.
XX
OS   Borrelia burgdorferi H9.
XX
PN   WO9425596-A.
XX
PD   10-NOV-1994.
XX
PF   29-APR-1994;   94WO-EP01365.
XX
PR   29-APR-1993;   93US-0053863.
XX
PA   (IMMO ) IMMUNO AG.
XX
PI   Crowe B,  Dorner F,  Livey I;
XX
DR   WPI; 1994-358273/44.
DR   N-PSDB; AAQ60894.
XX
PT   Immunogenic composition comprising OspC antigens - for the
PT   treatment of Lyme borreliosis in different, specific geographical
PT   areas.
XX
PS   Disclosure; Fig. 9a; 115pp; English.
XX
CC   A vaccine for Lyme disease includes selected OspC antigen
CC   formulations based on defined OspC families resolved by serovar
CC   typing and RFLP typing.  Partial sequences of OspC genes selected
CC   from different RFLP types are given in AAQ73883-905 (encoded peptides,
CC   comprising the first 92% of mature OspC, are given in AAR62771-93).
CC   Complete sequences of these novel ospC genes, including the 3' end,
CC   plus sequences for the ospC genes of Borrelia strains H13 and 28691
CC   are given in AAQ73857-82, and encoded proteins in AAR60884-909.  The
CC   DNA sequences may be expressed in e.g. Pichia pastoris for
CC   recombinant antigen production.
XX
SQ   Sequence   194 AA;


  Query Match             100.0%;  Score 52;  DB 15;  Length 194;
  Best Local Similarity   100.0%;  Pred. No. 0.15;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 PVVAESPKKP 10
          ||||||||||
Db    185 pvvaespkkp 194


RESULT  13
AAR60896
ID   AAR60896 standard; Protein; 194 AA.
XX
AC   AAR60896;
XX
DT   25-MAY-1995  (first entry)
XX
DE   Borrelia JSB antigen vaccine.
XX
KW   OspC antigen; vaccine; Lyme disease; borreliosis; immunogen;
KW   serovar typing; restriction fragment length polymorphism;
KW   RFLP; Pichia pastoris.
XX
OS   Borrelia burgdorferi JSB.
XX
PN   WO9425596-A.
XX
PD   10-NOV-1994.
XX
PF   29-APR-1994;   94WO-EP01365.
XX
PR   29-APR-1993;   93US-0053863.
XX
PA   (IMMO ) IMMUNO AG.
XX
PI   Crowe B,  Dorner F,  Livey I;
XX
DR   WPI; 1994-358273/44.
DR   N-PSDB; AAQ73869.
XX
PT   Immunogenic composition comprising OspC antigens - for the
PT   treatment of Lyme borreliosis in different, specific geographical
PT   areas.
XX
PS   Disclosure; Fig. 9a; 115pp; English.
XX
CC   A vaccine for Lyme disease includes selected OspC antigen
CC   formulations based on defined OspC families resolved by serovar
CC   typing and RFLP typing.  Partial sequences of OspC genes selected
CC   from different RFLP types are given in AAQ73883-905 (encoded peptides,
CC   comprising the first 92% of mature OspC, are given in AAR62771-93).
CC   Complete sequences of these novel ospC genes, including the 3' end,
CC   plus sequences for the ospC genes of Borrelia strains H13 and 28691
CC   are given in AAQ73857-82, and encoded proteins in AAR60884-909.  The
CC   DNA sequences may be expressed in e.g. Pichia pastoris for
CC   recombinant antigen production.
XX
SQ   Sequence   194 AA;


  Query Match             100.0%;  Score 52;  DB 15;  Length 194;
  Best Local Similarity   100.0%;  Pred. No. 0.15;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 PVVAESPKKP 10
          ||||||||||
Db    185 pvvaespkkp 194


RESULT  14
AAY78908
ID   AAY78908 standard; Protein; 194 AA.
XX
AC   AAY78908;
XX
DT   19-MAY-2000  (first entry)
XX
DE   Outer surface protein C (OspC) DraI fragment amino acid sequence.
XX
KW   Outer surface protein C; OspC; immunological epitope; Lyme disease;
KW   vaccine; prevention; Borrelia infection; diagnose.
XX
OS   Borrelia burgdorferi.
XX
PN   WO200006745-A1.
XX
PD   10-FEB-2000.
XX
PF   30-JUL-1999;   99WO-US17270.
XX
PR   31-JUL-1998;   98US-0094955.
```

XX
PA (GUND-) GUNDERSEN LUTHERAN MEDICAL FOUND INC.
XX
PI Callister SN, Lovrich SD, Schell RF, Jobe DA;
XX
DR WPI; 2000-195305/17.
DR N-PSDB; AAZ92216.
XX
PT New immunogenic polypeptides useful as a vaccine against Lyme disease
PT and for treating and detecting borrelia infection in mammals consists
PT an epitope of Borrelia burgdorferi OspC fragment -
XX
PS Claim 3; Fig 4; 51pp; English.
XX
CC This sequence represents the Borrelia burgdorferi outer surface protein C
CC (OspC) DraI fragment amino acid sequence. The polypeptide contains an
CC immunological epitope used in the invention. Large amounts of OspC are
CC rapidly synthesised by B. burgdorferi shortly after attachment of
CC infected ticks to mammalian hosts. The OspC protein sequence is used to
CC diagnose B. borrelia infection in mammals. The OspC nucleotide sequence
CC is used to prevent (via vaccination), treat or detect Borrelia
CC (especially B. burgdorferi) infections, i.e. Lyme disease, in mammals
CC including humans. The OspC nucleotide sequence provides a superior
CC diagnostic antigen that detects early Lyme disease infection, predicts
CC successful eradication or the organism from the host, and discriminates
CC between individuals with Lyme disease and individuals who have been
CC vaccinated with an OspA Lyme disease vaccination. Detection of anti-OspC
CC borreliacidal antibodies advantageously gives an early diagnosis which
CC anti-OspA and anti-OspB borreliacidal antibodies cannot do. Unlike
CC vaccination with OspA, vaccination with OspC results in clearance of
CC spirochetes and resolution of symptoms even if administered after
CC infection with B. burgdorferi.
XX
SQ Sequence 194 AA;

Query Match 100.0%; Score 52; DB 21; Length 194;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PVVAESPKKP 10
||||||||||
Db 185 pvvaespkkp 194

RESULT 15
AAB62723
ID AAB62723 standard; Protein; 206 AA.
XX
AC AAB62723;
XX
DT 03-APR-2001 (first entry)
XX
DE B afzelii ospC protein SEQ ID NO: 50.
XX
KW Borrelia; ospC; Lyme disease; vaccine; chimeric protein; tick.
XX
OS Borrelia afzelii.
XX
PN WO200078966-A1.
XX
PD 28-DEC-2000.
XX
PF 19-JUN-2000; 2000WO-US16915.
XX
PR 18-JUN-1999; 99US-0140042.
XX
PA (UYNY) UNIV NEW YORK STATE RES FOUND.
PA (BROO-) BROOK BIOTECHNOLOGIES INC.
XX
PI Dattwyler RJ, Seinost G, Dykhuizen D, Luft BJ, Gomes-Solecki M;
XX
DR WPI; 2001-050113/06.
DR N-PSDB; AAF29027.
XX
PT Compositions of OspC polypeptides from strains of Borrelia which cause
PT Lyme disease are used to immunize animals and detect immune responses
PT to Lyme disease -
XX
PS Disclosure; Page 109; 160pp; English.
XX
CC The present invention provides compositions comprising ospC proteins and
CC chimeric ospC proteins from members of the Borrelia genus. These may be
CC Borrelia burgdorferi, B. afzelii or B. garinii. These can be used as
CC vaccines against Borrelia infection, which is spread by ticks and leads
CC to Lyme disease.
XX
SQ Sequence 206 AA;

Query Match 100.0%; Score 52; DB 22; Length 206;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PVVAESPKKP 10
||||||||||
Db 197 pvvaespkkp 206

Search completed: October 12, 2002, 20:43:41
Job time: 13936 sec

GenCore version 4.5

OM protein - protein search, using sw model

Run on: October 12, 2002, 16:54:45 ; Search time 32.92 Seconds
(without alignments)
7.420 Million cell updates/sec

Title: US-09-408-578A-1
Perfect score: 52
Sequence: 1 PVVAESPKKP 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	52	100.0	194	4	US-09-364-083-2	Sequence 2, Appli
2	52	100.0	207	4	US-08-235-836C-36	Sequence 36, Appl
3	52	100.0	209	4	US-08-235-836C-32	Sequence 32, Appl
4	52	100.0	210	1	US-08-158-353-3	Sequence 3, Appli
5	52	100.0	210	4	US-08-209-603E-15	Sequence 15, Appl
6	52	100.0	210	4	US-08-235-836C-30	Sequence 30, Appl
7	52	100.0	212	1	US-08-158-353-4	Sequence 4, Appli
8	52	100.0	212	4	US-09-196-293-11	Sequence 11, Appl
9	52	100.0	212	4	US-08-209-603E-11	Sequence 11, Appl
10	52	100.0	212	4	US-08-235-836C-34	Sequence 34, Appl
11	52	100.0	466	4	US-08-235-836C-107	Sequence 107, App
12	52	100.0	466	4	US-08-235-836C-110	Sequence 110, App
13	52	100.0	588	4	US-08-235-836C-122	Sequence 122, App
14	49	94.2	212	1	US-08-031-295-2	Sequence 2, Appli
15	49	94.2	212	1	US-08-158-353-2	Sequence 2, Appli
16	49	94.2	212	4	US-07-903-580-2	Sequence 2, Appli
17	40	76.9	209	4	US-09-196-293-15	Sequence 15, Appl
18	35	67.3	1255	3	US-08-947-823-3	Sequence 3, Appli
19	35	67.3	1257	3	US-08-947-823-5	Sequence 5, Appli
20	34	65.4	984	1	US-08-242-932-2	Sequence 2, Appli
21	34	65.4	984	1	US-08-714-481-2	Sequence 2, Appli
22	34	65.4	984	5	PCT-US95-06111-2	Sequence 2, Appli
23	34	65.4	1098	4	US-08-923-992A-8	Sequence 8, Appli
24	34	65.4	1104	4	US-08-923-992A-4	Sequence 4, Appli
25	34	65.4	1128	4	US-08-923-992A-6	Sequence 6, Appli
26	34	65.4	1164	4	US-08-923-992A-2	Sequence 2, Appli
27	34	65.4	1164	4	US-08-923-992A-10	Sequence 10, Appl
28	33	63.5	22	4	US-08-557-006C-34	Sequence 34, Appl
29	33	63.5	197	3	US-08-415-655-6	Sequence 6, Appli
30	33	63.5	417	3	US-08-705-771-18	Sequence 18, Appl
31	33	63.5	492	4	US-09-342-749-2	Sequence 2, Appli
32	33	63.5	552	4	US-08-557-006C-40	Sequence 40, Appl
33	33	63.5	593	4	US-09-000-145-4	Sequence 4, Appli
34	33	63.5	933	3	US-08-293-728-2	Sequence 2, Appli
35	33	63.5	933	4	US-09-421-868-2	Sequence 2, Appli
36	32	61.5	164	1	US-08-357-125-4	Sequence 4, Appli
37	32	61.5	165	5	PCT-US95-03866-4	Sequence 4, Appli
38	32	61.5	273	1	US-08-220-379B-6	Sequence 6, Appli
39	32	61.5	273	1	US-08-341-456A-11	Sequence 11, Appl
40	32	61.5	273	2	US-08-478-414A-11	Sequence 11, Appl
41	32	61.5	273	3	US-08-325-240A-11	Sequence 11, Appl
42	32	61.5	273	4	US-08-898-982-11	Sequence 11, Appl
43	32	61.5	273	4	US-08-482-918-55	Sequence 55, Appl
44	32	61.5	273	4	US-09-224-681-55	Sequence 55, Appl
45	32	61.5	273	4	US-08-336-728A-55	Sequence 55, Appl

ALIGNMENTS

```
RESULT   1
US-09-364-083-2
; Sequence 2, Application US/09364083
; Patent No. 6210676
; GENERAL INFORMATION:
;  APPLICANT: Callister, Steven M
;  APPLICANT: Lovrich, Steven D
;  APPLICANT: Schell, Ronald F
;  APPLICANT: Jobe, Dean A
;  TITLE OF INVENTION: Compositions and Method Using the Borreliacidal
;  TITLE OF INVENTION: Epitope(s) of Borrelia Burgdorferi Outer Surface
;  TITLE OF INVENTION: Protein C (OspC) for the Diagnosis and Prevention of
;  TITLE OF INVENTION: Lyme Disease
;  FILE REFERENCE: B. burgdorferi  OspC
;  CURRENT APPLICATION NUMBER: US/09/364,083
;  CURRENT FILING DATE: 1999-07-30
;  EARLIER APPLICATION NUMBER: 60/094,955
;  EARLIER FILING DATE: 1998-07-31
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 194
;   TYPE: PRT
;   ORGANISM: Borrelia burgdorferi
US-09-364-083-2


  Query Match             100.0%;  Score 52;  DB 4;  Length 194;
  Best Local Similarity   100.0%;  Pred. No. 0.029;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 PVVAESPKKP 10
             ||||||||||
Db     185 PVVAESPKKP 194


RESULT   2
US-08-235-836C-36
; Sequence 36, Application US/08235836C
; Patent No. 6248562
;  GENERAL INFORMATION:
;    APPLICANT:  Dunn, John J.
;    APPLICANT:  Luft, Benjamin J.
;    TITLE OF INVENTION:  No. 6248562el Chimeric Proteins Comprising
;    TITLE OF INVENTION:  Borrelia Polypeptides and Uses Therefor
;    NUMBER OF SEQUENCES:  144
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Brookhaven National Laboratory
;      STREET:
;      CITY:  Upton
```

```
;       STATE:  NY
;       COUNTRY:  USA
;       ZIP:  11973
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/235,836C
;       FILING DATE:  29-APR-1994
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/148,191
;       FILING DATE:  01-11-93
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Bogosian, Margaret C.
;       REGISTRATION NUMBER:  25,324
;       REFERENCE/DOCKET NUMBER:  BNL93-28A
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (516) 282-7338
;       TELEFAX:  (516) 282-3729
;   INFORMATION FOR SEQ ID NO:  36:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  207 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-235-836C-36


  Query Match             100.0%;  Score 52;  DB 4;  Length 207;
  Best Local Similarity   100.0%;  Pred. No. 0.031;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 PVVAESPKKP 10
           ||||||||||
Db     198 PVVAESPKKP 207


RESULT   3
US-08-235-836C-32
; Sequence 32, Application US/08235836C
; Patent No. 6248562
;  GENERAL INFORMATION:
;    APPLICANT:  Dunn, John J.
;    APPLICANT:  Luft, Benjamin J.
;    TITLE OF INVENTION:  No. 6248562el Chimeric Proteins Comprising
;    TITLE OF INVENTION:  Borrelia Polypeptides and Uses Therefor
;    NUMBER OF SEQUENCES:  144
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Brookhaven National Laboratory
;      STREET:
;      CITY:  Upton
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  11973
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/235,836C
;      FILING DATE:  29-APR-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/148,191
;      FILING DATE:  01-11-93
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bogosian, Margaret C.
;      REGISTRATION NUMBER:  25,324
;      REFERENCE/DOCKET NUMBER:  BNL93-28A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (516) 282-7338
;      TELEFAX:  (516) 282-3729
;  INFORMATION FOR SEQ ID NO:  32:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  209 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-235-836C-32


  Query Match             100.0%;  Score 52;  DB 4;  Length 209;
  Best Local Similarity   100.0%;  Pred. No. 0.032;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 PVVAESPKKP 10
           ||||||||||
Db     200 PVVAESPKKP 209


RESULT   4
US-08-158-353-3
; Sequence 3, Application US/08158353
; Patent No. 5620862
;  GENERAL INFORMATION:
;    APPLICANT:  Padula, Steven J.
;    TITLE OF INVENTION:  Methods for Diagnosing Early Lyme
;    TITLE OF INVENTION:  Disease
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Hamilton, Brook, Smith & Reynolds, P.C.
;      STREET:  Two Militia Drive
;      CITY:  Lexington
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02173
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/158,353
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Carroll, Alice O.
;      REGISTRATION NUMBER:  33,542
;      REFERENCE/DOCKET NUMBER:  UCT93-05
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-861-6240
;      TELEFAX:  617-861-9540
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  210 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-158-353-3


  Query Match             100.0%;  Score 52;  DB 1;  Length 210;
  Best Local Similarity   100.0%;  Pred. No. 0.032;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 PVVAESPKKP 10
           ||||||||||
Db     201 PVVAESPKKP 210


RESULT   5
```

```
US-08-209-603E-15
; Sequence 15, Application US/08209603E
; Patent No. 6248538
;  GENERAL INFORMATION:
;    APPLICANT:  FUCHS, RENATE
;    APPLICANT:  WILSKE, BETTINA
;    APPLICANT:  PREAC-MURSIC, VERA
;    APPLICANT:  MOTZ, MANFRED
;    APPLICANT:  SOUTSCHECK, ERWIN
;    TITLE OF INVENTION:  IMMUNOLOGICALLY ACTIVE PROTEINS
;    TITLE OF INVENTION:  FROM BORRELIA BURGDORFERI
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BROOKS HAIDT HAFFNER & DELAHUNTY
;      STREET:  99 PARK AVENUE
;      CITY:  NEW YORK
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10016
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" FLOPPY DISC
;      COMPUTER:  AT&T - IBM COMPATIBLE
;      OPERATING SYSTEM:  MS-DOS Version 6.2
;      SOFTWARE:  ASCII
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/209,603E
;      FILING DATE:  10-MAR-1994
;      CLASSIFICATION:  436
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/EP90/02282
;      FILING DATE:  21-DEC-1990
;      APPLICATION NUMBER:  US 07/862,535
;      FILING DATE:  19-JUN-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  ROBINSON, WILLIAM R.
;      REGISTRATION NUMBER:  27,224
;      REFERENCE/DOCKET NUMBER:  LKR-9217-A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 697-3355
;      TELEFAX:  (212) 557-5635
;  INFORMATION FOR SEQ ID NO:  15:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  210
;      TYPE:  AMINO ACID
;      TOPOLOGY:  LINEAR
;    MOLECULE TYPE:
;      DESCRIPTION:  PROTEIN
;    HYPOTHETICAL:  N/A
;    ANTI-SENSE:  N/A
;    FRAGMENT TYPE:  N/A
;    ORIGINAL SOURCE:
;      ORGANISM:  B. BURGDORFERI
;    IMMEDIATE SOURCE:
;      LIBRARY:  DSM 5662
;    POSITION IN GENOME:  N/A
;    FEATURE:
;      IDENTIFICATION METHOD:  amino acid analysis
;    PUBLICATION INFORMATION:  N/A
US-08-209-603E-15


  Query Match             100.0%;  Score 52;  DB 4;  Length 210;
  Best Local Similarity   100.0%;  Pred. No. 0.032;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 PVVAESPKKP 10
           ||||||||||
Db     201 PVVAESPKKP 210


RESULT   6
US-08-235-836C-30
; Sequence 30, Application US/08235836C
; Patent No. 6248562
;  GENERAL INFORMATION:
;    APPLICANT:  Dunn, John J.
;    APPLICANT:  Luft, Benjamin J.
;    TITLE OF INVENTION:  No. 6248562el Chimeric Proteins Comprising
;    TITLE OF INVENTION:  Borrelia Polypeptides and Uses Therefor
;    NUMBER OF SEQUENCES:  144
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Brookhaven National Laboratory
;      STREET:
;      CITY:  Upton
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  11973
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/235,836C
;      FILING DATE:  29-APR-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/148,191
;      FILING DATE:  01-11-93
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bogosian, Margaret C.
;      REGISTRATION NUMBER:  25,324
;      REFERENCE/DOCKET NUMBER:  BNL93-28A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (516) 282-7338
;      TELEFAX:  (516) 282-3729
;  INFORMATION FOR SEQ ID NO:  30:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  210 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-235-836C-30


  Query Match             100.0%;  Score 52;  DB 4;  Length 210;
  Best Local Similarity   100.0%;  Pred. No. 0.032;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 PVVAESPKKP 10
           ||||||||||
Db     201 PVVAESPKKP 210


RESULT   7
US-08-158-353-4
; Sequence 4, Application US/08158353
; Patent No. 5620862
;  GENERAL INFORMATION:
;    APPLICANT:  Padula, Steven J.
;    TITLE OF INVENTION:  Methods for Diagnosing Early Lyme
;    TITLE OF INVENTION:  Disease
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Hamilton, Brook, Smith & Reynolds, P.C.
;      STREET:  Two Militia Drive
;      CITY:  Lexington
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02173
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
```

```
;       APPLICATION NUMBER:  US/08/158,353
;       FILING DATE:
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Carroll, Alice O.
;       REGISTRATION NUMBER:  33,542
;       REFERENCE/DOCKET NUMBER:  UCT93-05
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  617-861-6240
;       TELEFAX:  617-861-9540
;   INFORMATION FOR SEQ ID NO:  4:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  212 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-158-353-4


  Query Match             100.0%;  Score 52;  DB 1;  Length 212;
  Best Local Similarity   100.0%;  Pred. No. 0.032;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 PVVAESPKKP 10
           ||||||||||
Db     203 PVVAESPKKP 212


RESULT   8
US-09-196-293-11
; Sequence 11, Application US/09196293
; Patent No. 6183755
; GENERAL INFORMATION:
;  APPLICANT: Fuchs, Renate
;  APPLICANT: Motz, Manfred
;  APPLICANT: Soutscheck, Erwin
;  APPLICANT: Wilske, Bettina
;  APPLICANT: Preac-Mursic, Vera
;  TITLE OF INVENTION: Active proteins from Borrelia
;  TITLE OF INVENTION: burgdorferi
;  FILE REFERENCE: 738.001US2
;  CURRENT APPLICATION NUMBER: US/09/196,293
;  CURRENT FILING DATE: 1998-11-19
;  EARLIER APPLICATION NUMBER: US 08/209,603
;  EARLIER FILING DATE: 1994-03-10
;  EARLIER APPLICATION NUMBER: US 07/862,535
;  EARLIER FILING DATE: 1992-06-19
;  EARLIER APPLICATION NUMBER: WO PCT/EP90/02282
;  EARLIER FILING DATE: 1990-12-21
;  EARLIER APPLICATION NUMBER: DE P39 42 728.5
;  EARLIER FILING DATE: 1989-12-22
;  EARLIER APPLICATION NUMBER: DE P40 18 988.0
;  EARLIER FILING DATE: 1990-06-13
;  NUMBER OF SEQ ID NOS: 16
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11
;   LENGTH: 212
;   TYPE: PRT
;   ORGANISM: Borrelia burgdorferi
US-09-196-293-11


  Query Match             100.0%;  Score 52;  DB 4;  Length 212;
  Best Local Similarity   100.0%;  Pred. No. 0.032;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 PVVAESPKKP 10
           ||||||||||
Db     203 PVVAESPKKP 212


RESULT   9
US-08-209-603E-11
; Sequence 11, Application US/08209603E
; Patent No. 6248538
;  GENERAL INFORMATION:
;    APPLICANT:  FUCHS, RENATE
;    APPLICANT:  WILSKE, BETTINA
;    APPLICANT:  PREAC-MURSIC, VERA
;    APPLICANT:  MOTZ, MANFRED
;    APPLICANT:  SOUTSCHECK, ERWIN
;    TITLE OF INVENTION:  IMMUNOLOGICALLY ACTIVE PROTEINS
;    TITLE OF INVENTION:  FROM BORRELIA BURGDORFERI
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BROOKS HAIDT HAFFNER & DELAHUNTY
;      STREET:  99 PARK AVENUE
;      CITY:  NEW YORK
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10016
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" FLOPPY DISC
;      COMPUTER:  AT&T - IBM COMPATIBLE
;      OPERATING SYSTEM:  MS-DOS Version 6.2
;      SOFTWARE:  ASCII
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/209,603E
;      FILING DATE:  10-MAR-1994
;      CLASSIFICATION:  436
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/EP90/02282
;      FILING DATE:  21-DEC-1990
;      APPLICATION NUMBER:  US 07/862,535
;      FILING DATE:  19-JUN-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  ROBINSON, WILLIAM R.
;      REGISTRATION NUMBER:  27,224
;      REFERENCE/DOCKET NUMBER:  LKR-9217-A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 697-3355
;      TELEFAX:  (212) 557-5635
;  INFORMATION FOR SEQ ID NO:  11:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  212
;      TYPE:  AMINO ACID
;      TOPOLOGY:  LINEAR
;    MOLECULE TYPE:
;      DESCRIPTION:  PROTEIN
;    HYPOTHETICAL:  N/A
;    ANTI-SENSE:  N/A
;    FRAGMENT TYPE:  N/A
;    ORIGINAL SOURCE:
;      ORGANISM:  B. BURGDORFERI
;    IMMEDIATE SOURCE:
;      LIBRARY:  DSM 5662
;    POSITION IN GENOME:  N/A
;    FEATURE:
;      IDENTIFICATION METHOD:  amino acid analysis
;    PUBLICATION INFORMATION:  N/A
;    ANTI-SENSE:  N/A
;    FRAGMENT TYPE:  INTERNAL
;    ORIGINAL SOURCE:
US-08-209-603E-11


  Query Match             100.0%;  Score 52;  DB 4;  Length 212;
  Best Local Similarity   100.0%;  Pred. No. 0.032;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 PVVAESPKKP 10
           ||||||||||
Db     203 PVVAESPKKP 212
```

```
RESULT  10
US-08-235-836C-34
; Sequence 34, Application US/08235836C
; Patent No. 6248562
;  GENERAL INFORMATION:
;    APPLICANT:  Dunn, John J.
;    APPLICANT:  Luft, Benjamin J.
;    TITLE OF INVENTION:  No. 6248562el Chimeric Proteins Comprising
;    TITLE OF INVENTION:  Borrelia Polypeptides and Uses Therefor
;    NUMBER OF SEQUENCES:  144
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Brookhaven National Laboratory
;      STREET:
;      CITY:  Upton
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  11973
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/235,836C
;      FILING DATE:  29-APR-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/148,191
;      FILING DATE:  01-11-93
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bogosian, Margaret C.
;      REGISTRATION NUMBER:  25,324
;      REFERENCE/DOCKET NUMBER:  BNL93-28A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (516) 282-7338
;      TELEFAX:  (516) 282-3729
;  INFORMATION FOR SEQ ID NO:  34:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  212 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-235-836C-34


  Query Match           100.0%;  Score 52;  DB 4;  Length 212;
  Best Local Similarity  100.0%;  Pred. No. 0.032;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 PVVAESPKKP 10
          ||||||||||
Db    203 PVVAESPKKP 212


RESULT  11
US-08-235-836C-107
; Sequence 107, Application US/08235836C
; Patent No. 6248562
;  GENERAL INFORMATION:
;    APPLICANT:  Dunn, John J.
;    APPLICANT:  Luft, Benjamin J.
;    TITLE OF INVENTION:  No. 6248562el Chimeric Proteins Comprising
;    TITLE OF INVENTION:  Borrelia Polypeptides and Uses Therefor
;    NUMBER OF SEQUENCES:  144
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Brookhaven National Laboratory
;      STREET:
;      CITY:  Upton
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  11973
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/235,836C
;      FILING DATE:  29-APR-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/148,191
;      FILING DATE:  01-11-93
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bogosian, Margaret C.
;      REGISTRATION NUMBER:  25,324
;      REFERENCE/DOCKET NUMBER:  BNL93-28A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (516) 282-7338
;      TELEFAX:  (516) 282-3729
;  INFORMATION FOR SEQ ID NO:  107:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  466 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-235-836C-107


  Query Match           100.0%;  Score 52;  DB 4;  Length 466;
  Best Local Similarity  100.0%;  Pred. No. 0.075;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 PVVAESPKKP 10
          ||||||||||
Db    457 PVVAESPKKP 466


RESULT  12
US-08-235-836C-110
; Sequence 110, Application US/08235836C
; Patent No. 6248562
;  GENERAL INFORMATION:
;    APPLICANT:  Dunn, John J.
;    APPLICANT:  Luft, Benjamin J.
;    TITLE OF INVENTION:  No. 6248562el Chimeric Proteins Comprising
;    TITLE OF INVENTION:  Borrelia Polypeptides and Uses Therefor
;    NUMBER OF SEQUENCES:  144
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Brookhaven National Laboratory
;      STREET:
;      CITY:  Upton
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  11973
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/235,836C
;      FILING DATE:  29-APR-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/148,191
;      FILING DATE:  01-11-93
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bogosian, Margaret C.
;      REGISTRATION NUMBER:  25,324
;      REFERENCE/DOCKET NUMBER:  BNL93-28A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (516) 282-7338
;      TELEFAX:  (516) 282-3729
;  INFORMATION FOR SEQ ID NO:  110:
;    SEQUENCE CHARACTERISTICS:
```

```
;      LENGTH:  466 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-235-836C-110


  Query Match             100.0%;  Score 52;  DB 4;  Length 466;
  Best Local Similarity   100.0%;  Pred. No. 0.075;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 PVVAESPKKP 10
          ||||||||||
Db    201 PVVAESPKKP 210


RESULT  13
US-08-235-836C-122
; Sequence 122, Application US/08235836C
; Patent No. 6248562
;  GENERAL INFORMATION:
;    APPLICANT:  Dunn, John J.
;    APPLICANT:  Luft, Benjamin J.
;    TITLE OF INVENTION:  No. 6248562el Chimeric Proteins Comprising
;    TITLE OF INVENTION:  Borrelia Polypeptides and Uses Therefor
;    NUMBER OF SEQUENCES:  144
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Brookhaven National Laboratory
;      STREET:
;      CITY:  Upton
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  11973
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/235,836C
;      FILING DATE:  29-APR-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/148,191
;      FILING DATE:  01-11-93
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bogosian, Margaret C.
;      REGISTRATION NUMBER:  25,324
;      REFERENCE/DOCKET NUMBER:  BNL93-28A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (516) 282-7338
;      TELEFAX:  (516) 282-3729
;  INFORMATION FOR SEQ ID NO:  122:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  588 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-235-836C-122


  Query Match             100.0%;  Score 52;  DB 4;  Length 588;
  Best Local Similarity   100.0%;  Pred. No. 0.097;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 PVVAESPKKP 10
          ||||||||||
Db    467 PVVAESPKKP 476


RESULT  14
US-08-031-295-2
; Sequence 2, Application US/08031295
; Patent No. 5530103
;  GENERAL INFORMATION:
;    APPLICANT:  LIVEY, Ian
;    APPLICANT:  DORNER, Freidrich
;    TITLE OF INVENTION:  METHOD AND COMPOSITION FOR THE
;    TITLE OF INVENTION:  PREVENTION OF LYME DISEASE
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  3000 K Street, N.W., Suite 500
;      CITY:  Washington, D.C.
;      COUNTRY:  USA
;      ZIP:  20007-5109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/031,295
;      FILING DATE:  19930312
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/903,580
;      FILING DATE:  25-JUN-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/824,161
;      FILING DATE:  22-JAN-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/727,245
;      FILING DATE:  11-JUL-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BENT, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  30472/142 IMMU
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)672-5300
;      TELEFAX:  (202)672-5399
;      TELEX:  904136
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  212 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-031-295-2


  Query Match             94.2%;  Score 49;  DB 1;  Length 212;
  Best Local Similarity   90.0%;  Pred. No. 0.11;
  Matches    9;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 PVVAESPKKP 10
          |||||:||||
Db    203 PVVAETPKKP 212


RESULT  15
US-08-158-353-2
; Sequence 2, Application US/08158353
; Patent No. 5620862
;  GENERAL INFORMATION:
;    APPLICANT:  Padula, Steven J.
;    TITLE OF INVENTION:  Methods for Diagnosing Early Lyme
;    TITLE OF INVENTION:  Disease
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Hamilton, Brook, Smith & Reynolds, P.C.
;      STREET:  Two Militia Drive
;      CITY:  Lexington
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02173
```

```
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/158,353
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Carroll, Alice O.
;      REGISTRATION NUMBER:  33,542
;      REFERENCE/DOCKET NUMBER:  UCT93-05
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-861-6240
;      TELEFAX:  617-861-9540
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  212 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-158-353-2


  Query Match             94.2%;  Score 49;  DB 1;  Length 212;
  Best Local Similarity   90.0%;  Pred. No. 0.11;
  Matches    9;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 PVVAESPKKP 10
           |||||:||||
Db     203 PVVAENPKKP 212


Search completed: October 12, 2002, 20:44:28
Job time: 13783 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: October 12, 2002, 17:08:10 ; Search time 49.13 Seconds
(without alignments)
19.558 Million cell updates/sec

Title: US-09-408-578A-1
Perfect score: 52
Sequence: 1 PVVAESPKKP 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	52	100.0	190	2	S70273	outer surface prot
2	52	100.0	191	2	S70278	outer surface prot
3	52	100.0	191	2	S70284	outer surface prot
4	52	100.0	193	2	S70265	outer surface prot
5	52	100.0	193	2	S70274	outer surface prot
6	52	100.0	193	2	S70279	outer surface prot
7	52	100.0	194	2	S70268	outer surface prot
8	52	100.0	194	2	S70277	outer surface prot
9	52	100.0	207	2	I40271	outer surface prot
10	52	100.0	207	2	I40276	outer surface prot
11	52	100.0	207	2	S69919	outer surface prot
12	52	100.0	207	2	S69924	outer surface prot
13	52	100.0	207	2	S37727	outer surface prot
14	52	100.0	209	2	I40273	outer surface prot
15	52	100.0	209	2	I40270	outer surface prot
16	52	100.0	209	2	S69926	outer surface prot
17	52	100.0	209	2	I40281	outer surface prot
18	52	100.0	209	2	I40285	outer surface prot
19	52	100.0	209	2	S69917	outer surface prot
20	52	100.0	210	2	I40272	outer surface prot
21	52	100.0	210	2	S69920	outer surface prot
22	52	100.0	210	2	S69925	outer surface prot
23	52	100.0	210	2	I40280	outer surface prot
24	52	100.0	210	2	I40284	outer surface prot
25	52	100.0	210	2	G70218	outer surface prot
26	52	100.0	210	2	I40144	outer surface prot
27	52	100.0	210	2	S69927	outer surface prot
28	52	100.0	210	2	S69923	outer surface prot
29	52	100.0	211	2	I40277	outer surface prot
30	52	100.0	211	2	I40278	outer surface prot
31	52	100.0	211	2	I40282	outer surface prot
32	52	100.0	211	2	I40145	outer surface prot
33	52	100.0	211	2	I40268	outer surface prot
34	52	100.0	211	2	S69930	outer surface prot
35	52	100.0	211	2	S69918	outer surface prot
36	52	100.0	211	2	S69932	outer surface prot
37	52	100.0	211	2	S69928	outer surface prot
38	52	100.0	211	2	S69929	outer surface prot
39	52	100.0	212	2	I40279	outer surface prot
40	52	100.0	212	2	S69921	outer surface prot
41	52	100.0	212	2	S69922	outer surface prot
42	52	100.0	212	2	I40275	outer surface prot
43	52	100.0	212	2	S20543	outer surface prot
44	52	100.0	214	2	S69916	outer surface prot
45	51	98.1	191	2	I40153	outer surface prot

ALIGNMENTS

RESULT 1
S70273
outer surface protein C - Lyme disease spirochete
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
C;Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 26-May-2000
C;Accession: S70273
R;Livey, I.; Gibbs, C.P.; Schuster, R.; Dorner, F.
Mol. Microbiol. 18, 257-269, 1995
A;Title: Evidence for lateral transfer and recombination in OspC variation in Lyme d
A;Reference number: S70255; MUID:96296448
A;Accession: S70273
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-190 <LIV>
A;Cross-references: EMBL:L42870; NID:g858737; PIDN:AAB37013.1; PID:g1695228
A;Experimental source: strain VSDA
C;Genetics:
A;Gene: ospC
C;Superfamily: Lyme disease spirochete surface protein C

Query Match 100.0%; Score 52; DB 2; Length 190;
Best Local Similarity 100.0%; Pred. No. 0.036;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PVVAESPKKP 10
||||||||||
Db 181 PVVAESPKKP 190

RESULT 2
S70278
outer surface protein C - Lyme disease spirochete
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
C;Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 26-May-2000
C;Accession: S70278
R;Livey, I.; Gibbs, C.P.; Schuster, R.; Dorner, F.
Mol. Microbiol. 18, 257-269, 1995
A;Title: Evidence for lateral transfer and recombination in OspC variation in Lyme d
A;Reference number: S70255; MUID:96296448
A;Accession: S70278
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-191 <LIV>
A;Cross-references: EMBL:L42871; NID:g858738; PIDN:AAB37014.1; PID:g1695229
A;Experimental source: strain VS461
C;Genetics:
A;Gene: ospC
C;Superfamily: Lyme disease spirochete surface protein C

Query Match 100.0%; Score 52; DB 2; Length 191;

```
  Best Local Similarity   100.0%;  Pred. No. 0.036;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 PVVAESPKKP 10
          ||||||||||
Db    182 PVVAESPKKP 191


RESULT   3
S70284
outer surface protein C - Lyme disease spirochete
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
C;Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 26-May-2000
C;Accession: S70284
R;Livey, I.; Gibbs, C.P.; Schuster, R.; Dorner, F.
Mol. Microbiol. 18, 257-269, 1995
A;Title: Evidence for lateral transfer and recombination in OspC variation in Lyme disea
A;Reference number: S70255; MUID:96296448
A;Accession: S70284
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-191 <LIV>
A;Cross-references: EMBL:L42896; NID:g858724; PIDN:AAB37004.1; PID:g1695221
A;Experimental source: strain 27579
C;Genetics:
A;Gene: ospC
C;Superfamily: Lyme disease spirochete surface protein C


  Query Match             100.0%;  Score 52;  DB 2;  Length 191;
  Best Local Similarity   100.0%;  Pred. No. 0.036;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 PVVAESPKKP 10
          ||||||||||
Db    182 PVVAESPKKP 191


RESULT   4
S70265
outer surface protein C - Lyme disease spirochete
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
C;Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 26-May-2000
C;Accession: S70265
R;Livey, I.; Gibbs, C.P.; Schuster, R.; Dorner, F.
Mol. Microbiol. 18, 257-269, 1995
A;Title: Evidence for lateral transfer and recombination in OspC variation in Lyme disea
A;Reference number: S70255; MUID:96296448
A;Accession: S70265
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-193 <LIV>
A;Cross-references: EMBL:L42884; NID:g858710; PIDN:AAB36992.1; PID:g1695210
A;Experimental source: strain J1
C;Genetics:
A;Gene: ospC
C;Superfamily: Lyme disease spirochete surface protein C


  Query Match             100.0%;  Score 52;  DB 2;  Length 193;
  Best Local Similarity   100.0%;  Pred. No. 0.036;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 PVVAESPKKP 10
          ||||||||||
Db    184 PVVAESPKKP 193


RESULT   5
S70274
outer surface protein C - Lyme disease spirochete
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
C;Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 26-May-2000
C;Accession: S70274
R;Livey, I.; Gibbs, C.P.; Schuster, R.; Dorner, F.
Mol. Microbiol. 18, 257-269, 1995
A;Title: Evidence for lateral transfer and recombination in OspC variation in Lyme
A;Reference number: S70255; MUID:96296448
A;Accession: S70274
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-193 <LIV>
A;Cross-references: EMBL:L42892; NID:g858720; PIDN:AAB37000.1; PID:g1695217
A;Experimental source: strain acal
C;Genetics:
A;Gene: ospC
C;Superfamily: Lyme disease spirochete surface protein C


  Query Match             100.0%;  Score 52;  DB 2;  Length 193;
  Best Local Similarity   100.0%;  Pred. No. 0.036;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 PVVAESPKKP 10
          ||||||||||
Db    184 PVVAESPKKP 193


RESULT   6
S70279
outer surface protein C - Lyme disease spirochete
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
C;Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 26-May-2000
C;Accession: S70279
R;Livey, I.; Gibbs, C.P.; Schuster, R.; Dorner, F.
Mol. Microbiol. 18, 257-269, 1995
A;Title: Evidence for lateral transfer and recombination in OspC variation in Lyme
A;Reference number: S70255; MUID:96296448
A;Accession: S70279
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-193 <LIV>
A;Cross-references: EMBL:L42898; NID:g858729; PIDN:AAB37007.1; PID:g1695223
A;Experimental source: strain 25015
C;Genetics:
A;Gene: ospC
C;Superfamily: Lyme disease spirochete surface protein C


  Query Match             100.0%;  Score 52;  DB 2;  Length 193;
  Best Local Similarity   100.0%;  Pred. No. 0.036;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 PVVAESPKKP 10
          ||||||||||
Db    184 PVVAESPKKP 193


RESULT   7
S70268
outer surface protein C - Lyme disease spirochete
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
C;Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 26-May-2000
C;Accession: S70268
R;Livey, I.; Gibbs, C.P.; Schuster, R.; Dorner, F.
Mol. Microbiol. 18, 257-269, 1995
A;Title: Evidence for lateral transfer and recombination in OspC variation in Lyme
A;Reference number: S70255; MUID:96296448
A;Accession: S70268
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-194 <LIV>
A;Cross-references: EMBL:L42888; NID:g858716; PIDN:AAB36996.1; PID:g1695213
A;Experimental source: strain H9
C;Genetics:
A;Gene: ospC
```

```
C;Superfamily: Lyme disease spirochete surface protein C


  Query Match            100.0%;  Score 52;  DB 2;  Length 194;
  Best Local Similarity  100.0%;  Pred. No. 0.037;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 PVVAESPKKP 10
          ||||||||||
Db    185 PVVAESPKKP 194


RESULT   8
S70277
outer surface protein C - Lyme disease spirochete
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
C;Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 26-May-2000
C;Accession: S70277
R;Livey, I.; Gibbs, C.P.; Schuster, R.; Dorner, F.
Mol. Microbiol. 18, 257-269, 1995
A;Title: Evidence for lateral transfer and recombination in OspC variation in Lyme disea
A;Reference number: S70255; MUID:96296448
A;Accession: S70277
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-194 <LIV>
A;Cross-references: EMBL:L42873; NID:g858740; PIDN:AAB37016.1; PID:g1695231
A;Experimental source: strain SIMON
C;Genetics:
A;Gene: ospC
C;Superfamily: Lyme disease spirochete surface protein C


  Query Match            100.0%;  Score 52;  DB 2;  Length 194;
  Best Local Similarity  100.0%;  Pred. No. 0.037;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 PVVAESPKKP 10
          ||||||||||
Db    185 PVVAESPKKP 194


RESULT   9
I40271
outer surface protein C precursor - Borrelia garinii
C;Species: Borrelia garinii
C;Date: 04-Sep-1997 #sequence_revision 04-Sep-1997 #text_change 26-May-2000
C;Accession: I40271
R;Fukunaga, M.; Hamase, A.
J. Clin. Microbiol. 33, 2415-2420, 1995
A;Title: Outer surface protein C gene sequence analysis of Borrelia burgdorferi sensu la
A;Reference number: I40269; MUID:96025162
A;Accession: I40271
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-207 <RES>
A;Cross-references: GB:D49377; NID:g1041103; PIDN:BAA08375.1; PID:g1041104
C;Superfamily: Lyme disease spirochete surface protein C


  Query Match            100.0%;  Score 52;  DB 2;  Length 207;
  Best Local Similarity  100.0%;  Pred. No. 0.039;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 PVVAESPKKP 10
          ||||||||||
Db    198 PVVAESPKKP 207


RESULT  10
I40276
outer surface protein C precursor - Borrelia garinii
C;Species: Borrelia garinii
C;Date: 04-Sep-1997 #sequence_revision 04-Sep-1997 #text_change 26-May-2000
C;Accession: I40276
R;Fukunaga, M.; Hamase, A.
J. Clin. Microbiol. 33, 2415-2420, 1995
A;Title: Outer surface protein C gene sequence analysis of Borrelia burgdorferi sens
A;Reference number: I40269; MUID:96025162
A;Accession: I40276
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-207 <RES>
A;Cross-references: GB:D49500; NID:g707095; PIDN:BAA08460.1; PID:g769687
C;Superfamily: Lyme disease spirochete surface protein C


  Query Match            100.0%;  Score 52;  DB 2;  Length 207;
  Best Local Similarity  100.0%;  Pred. No. 0.039;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 PVVAESPKKP 10
          ||||||||||
Db    198 PVVAESPKKP 207


RESULT  11
S69919
outer surface protein C precursor - Borrelia garinii (strain PTrob)
C;Species: Borrelia garinii
A;Variety: strain PTrob
C;Date: 06-Dec-1996 #sequence_revision 14-Feb-1997 #text_change 26-May-2000
C;Accession: S69919
R;Jauris-Heipke, S.; Liegl, G.; Preac-Mursic, V.; Roessler, D.; Schwab, E.; Soutsche
J. Clin. Microbiol. 33, 1860-1866, 1995
A;Title: Molecular analysis of genes encoding outer surface protein C (OspC) of Borr
A;Reference number: I40047; MUID:95395018
A;Accession: S69919
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-207 <JAU>
A;Cross-references: EMBL:X83554; NID:g872027; PIDN:CAA58544.1; PID:g872028
A;Experimental source: strain PTrob
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1995
C;Genetics:
A;Gene: ospC
C;Superfamily: Lyme disease spirochete surface protein C
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-207/Product: outer surface protein C #status predicted <MAT>


  Query Match            100.0%;  Score 52;  DB 2;  Length 207;
  Best Local Similarity  100.0%;  Pred. No. 0.039;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 PVVAESPKKP 10
          ||||||||||
Db    198 PVVAESPKKP 207


RESULT  12
S69924
outer surface protein C precursor - Borrelia garinii (strain TIs1)
C;Species: Borrelia garinii
A;Variety: strain TIs1
C;Date: 15-Feb-1997 #sequence_revision 27-Feb-1997 #text_change 26-May-2000
C;Accession: S69924; S72677
R;Jauris-Heipke, S.; Liegl, G.; Preac-Mursic, V.; Roessler, D.; Schwab, E.; Soutsche
J. Clin. Microbiol. 33, 1860-1866, 1995
A;Title: Molecular analysis of genes encoding outer surface protein C (OspC) of Borr
A;Reference number: I40047; MUID:95395018
A;Accession: S69924
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-207 <JAU>
A;Cross-references: EMBL:X81525
```

A;Experimental source: strain TIs1
R;Roessler, D.
submitted to the EMBL Data Library, September 1994
A;Reference number: S72674
A;Accession: S72677
A;Molecule type: DNA
A;Residues: 1-77,'VE',80-207 <ROE>
A;Cross-references: EMBL:X81525; NID:g804962; PIDN:CAA57245.1; PID:g804963
C;Genetics:
A;Gene: ospC
C;Superfamily: Lyme disease spirochete surface protein C

Query Match 100.0%; Score 52; DB 2; Length 207;
Best Local Similarity 100.0%; Pred. No. 0.039;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PVVAESPKKP 10
||||||||||
Db 198 PVVAESPKKP 207

RESULT 13
S37727
outer surface protein C precursor - Lyme disease spirochete
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 26-May-2000
C;Accession: S37727
R;Jauris-Heipke, S.; Fuchs, R.; Motz, M.; Preac-Mursic, V.; Schwab, E.; Soutschek, E.; W
Med. Microbiol. Immunol. 182, 37-50, 1993
A;Title: Genetic heterogenity of the genes coding for the outer surface protein C (OspC)
A;Reference number: S37726; MUID:93268136
A;Accession: S37727
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-207 <JAU>
A;Cross-references: EMBL:X69595; NID:g311393; PIDN:CAA49305.1; PID:g311394
C;Superfamily: Lyme disease spirochete surface protein C

Query Match 100.0%; Score 52; DB 2; Length 207;
Best Local Similarity 100.0%; Pred. No. 0.039;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PVVAESPKKP 10
||||||||||
Db 198 PVVAESPKKP 207

RESULT 14
I40273
outer surface protein C precursor - Borrelia afzelii
C;Species: Borrelia afzelii
C;Date: 04-Sep-1997 #sequence_revision 04-Sep-1997 #text_change 26-May-2000
C;Accession: I40273
R;Fukunaga, M.; Hamase, A.
J. Clin. Microbiol. 33, 2415-2420, 1995
A;Title: Outer surface protein C gene sequence analysis of Borrelia burgdorferi sensu la
A;Reference number: I40269; MUID:96025162
A;Accession: I40273
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-209 <RES>
A;Cross-references: GB:D49379; NID:g1041107; PIDN:BAA08377.1; PID:g1041108
C;Superfamily: Lyme disease spirochete surface protein C

Query Match 100.0%; Score 52; DB 2; Length 209;
Best Local Similarity 100.0%; Pred. No. 0.039;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PVVAESPKKP 10
||||||||||
Db 200 PVVAESPKKP 209

RESULT 15
I40270
outer surface protein C precursor - Borrelia garinii
C;Species: Borrelia garinii
C;Date: 04-Sep-1997 #sequence_revision 04-Sep-1997 #text_change 26-May-2000
C;Accession: I40270
R;Fukunaga, M.; Hamase, A.
J. Clin. Microbiol. 33, 2415-2420, 1995
A;Title: Outer surface protein C gene sequence analysis of Borrelia burgdorferi ser
A;Reference number: I40269; MUID:96025162
A;Accession: I40270
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-209 <RES>
A;Cross-references: GB:D49498; NID:g707093; PIDN:BAA08458.1; PID:g769685
C;Superfamily: Lyme disease spirochete surface protein C

Query Match 100.0%; Score 52; DB 2; Length 209;
Best Local Similarity 100.0%; Pred. No. 0.039;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PVVAESPKKP 10
||||||||||
Db 200 PVVAESPKKP 209

Search completed: October 12, 2002, 20:45:30
Job time: 13040 sec

GenCore version 4.5

OM protein - protein search, using sw model

Run on: October 12, 2002, 20:42:41 ; Search time 35.71 Seconds
(without alignments)
10.843 Million cell updates/sec

Title: US-09-408-578A-1
Perfect score: 52
Sequence: 1 PVVAESPKKP 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	52	100.0	210	1	OSC1_BORBU	Q07337 borrelia bu
2	52	100.0	212	1	OSC2_BORBU	Q08137 borrelia bu
3	37	71.2	205	1	CLP1_RHIME	P58277 rhizobium m
4	37	71.2	502	1	RP54_AZOVI	P08623 azotobacter
5	37	71.2	865	1	CPN_DROME	Q02910 drosophila
6	36	69.2	122	1	YRB1_SYNP6	P23655 synechococc
7	36	69.2	382	1	YJO7_YEAST	P47007 saccharomyc
8	36	69.2	2150	1	SDC3_CAEEL	P34706 caenorhabdi
9	35	67.3	297	1	HMX1_BOVIN	O02786 bos taurus
10	35	67.3	297	1	HMX1_HUMAN	P28360 homo sapien
11	35	67.3	299	1	ODC_HUMAN	Q9bqt8 homo sapien
12	35	67.3	1388	1	HRP3_SCHPO	O14139 schizosacch
13	34	65.4	102	1	VE7_PAPVD	P03131 deer papill
14	34	65.4	387	1	PAB_PEPMA	Q51911 peptostrept
15	34	65.4	430	1	SYTB_MOUSE	Q9r0n3 mus musculu
16	34	65.4	430	1	SYTB_RAT	O08835 rattus norv
17	34	65.4	431	1	SYTB_HUMAN	Q9bt88 homo sapien
18	34	65.4	482	1	DOK1_MOUSE	P97465 mus musculu
19	34	65.4	486	1	PRL1_ARATH	Q42384 arabidopsis
20	34	65.4	489	1	FLIF_AQUAE	O67241 aquifex aeo
21	34	65.4	528	1	ASM4_YEAST	Q05166 saccharomyc
22	34	65.4	599	1	TRM2_ARATH	Q9sru7 arabidopsis
23	34	65.4	627	1	DTNB_HUMAN	O60941 homo sapien
24	34	65.4	700	1	DTNB_MOUSE	O70585 mus musculu
25	34	65.4	1164	1	BAG_STRAG	P27951 streptococc
26	34	65.4	3707	1	PGBM_MOUSE	Q05793 mus musculu
27	33.5	64.4	216	1	H1C1_XENLA	P15866 xenopus lae
28	33.5	64.4	220	1	H1C2_XENLA	P15867 xenopus lae
29	33	63.5	209	1	H1A_XENLA	P06892 xenopus lae
30	33	63.5	211	1	ROPB_RHILV	Q52866 rhizobium l
31	33	63.5	219	1	H1B_XENLA	P06893 xenopus lae
32	33	63.5	288	1	HMXX_CHICK	P50223 gallus gall
33	33	63.5	333	1	GALE_STRMU	P96995 streptococc
34	33	63.5	372	1	TOLA_HAEIN	P44678 haemophilus
35	33	63.5	379	1	PSPB_DICDI	P54704 dictyosteli
36	33	63.5	388	1	MPT1_YEAST	P50105 saccharomyc
37	33	63.5	416	1	TRA1_HUMAN	Q13077 homo sapien
38	33	63.5	472	1	SX14_DROME	P40656 drosophila
39	33	63.5	492	1	TMS2_HUMAN	O15393 homo sapien
40	33	63.5	552	1	AAK2_HUMAN	P54646 homo sapien
41	33	63.5	552	1	AAK2_RAT	Q09137 rattus norv
42	33	63.5	587	1	NDC2_RAT	P70545 rattus norv
43	33	63.5	616	1	PRLR_RABIT	P14787 oryctolagus
44	33	63.5	660	1	TP6B_METJA	Q58434 methanococc
45	33	63.5	885	1	ASE1_YEAST	P50275 saccharomyc

ALIGNMENTS

RESULT 1
OSC1_BORBU
ID OSC1_BORBU STANDARD; PRT; 210 AA.
AC Q07337;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Outer surface protein C precursor (PC).
GN OSPC OR BBB19.
OS Borrelia burgdorferi (Lyme disease spirochete).
OG Plasmid lp54.
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35210 / B31;
RX MEDLINE=93268136; PubMed=8098841;
RA Jauris-Heipke S., Fuchs R., Motz M., Preac-Mursic V., Schwab E.,
RA Will G., Wilske B.;
RT "Genetic heterogenity of the genes coding for the outer surface
RT protein C (OspC) and the flagellin of Borrelia burgdorferi.";
RL Med. Microbiol. Immunol. 182:37-50(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35210 / B31;
RX MEDLINE=93239332; PubMed=8478108;
RA Wilske B., Preace-Mursic V., Jauris S., Pradel I., Soutschek E.,
RA Schwab E., Wanner G.;
RT "Immunological and molecular polymorphisms of OspC, an immunodominant
RT major outer surface protein of Borrelia burgdorferi.";
RL Infect. Immun. 61:2182-2191(1993).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35210 / B31;
RX MEDLINE=94041630; PubMed=8225587;
RA Padula S.J., Sampieri A., Dias F., Szczepanski A., Ryan R.W.;
RT "Molecular characterization and expression of p23 (OspC) from a North
RT American strain of Borrelia burgdorferi.";
RL Infect. Immun. 61:5097-5105(1993).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35210 / B31;
RX MEDLINE=96025162; PubMed=7494039;
RA Fukunaga M., Hamase A.;
RT "Outer surface protein C gene sequence analysis of Borrelia
RT burgdorferi sensu lato isolates from Japan.";
RL J. Clin. Microbiol. 33:2415-2420(1995).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35210 / B31;
RX MEDLINE=98065943; PubMed=9403685;
RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
RA Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,
RA Peterson J., Kerlavage A.R., Quackenbush J., Salzberg S., Hanson M.,
RA van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,

```
RA   Utterback T., Watthey L., McDonald L., Artiach P., Bowman C.,
RA   Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
RA   Smith H.O., Venter J.C.;
RT   "Genomic sequence of a Lyme disease spirochaete, Borrelia
RT   burgdorferi.";
RL   Nature 390:580-586(1997).
CC   -!- FUNCTION: NOT KNOWN; MAJOR IMMUNODOMINANT PROTEIN.
CC   -!- SUBCELLULAR LOCATION: Attached to the outer membrane by a lipid
CC       anchor.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X69596; CAA49306.1; -.
DR   EMBL; U01894; AAA16058.1; -.
DR   EMBL; D49497; BAA08457.1; -.
DR   EMBL; AE000792; AAC66329.1; -.
DR   TIGR; BBB19; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
DR   PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
KW   Outer membrane; Lipoprotein; Signal; Plasmid; Antigen;
KW   Complete proteome.
FT   SIGNAL        1     18       BY SIMILARITY.
FT   CHAIN        19    210       OUTER SURFACE PROTEIN C.
FT   LIPID        19     19       N-ACYL DIGLYCERIDE (BY SIMILARITY).
SQ   SEQUENCE   210 AA;  22340 MW;  7A4FC978F91777BF CRC64;


  Query Match             100.0%;  Score 52;  DB 1;  Length 210;
  Best Local Similarity   100.0%;  Pred. No. 0.028;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 PVVAESPKKP 10
          ||||||||||
Db    201 PVVAESPKKP 210


RESULT   2
OSC2_BORBU
ID   OSC2_BORBU     STANDARD;      PRT;   212 AA.
AC   Q08137;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   Outer surface protein C precursor (PC).
GN   OSPC.
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC   STRAIN=PKO;
RX   MEDLINE=92219995; PubMed=1560779;
RA   Fuchs R., Jauris S., Lottspeich F., Preac-Mursic V., Wilske B.,
RA   Soutschek E.;
RT   "Molecular analysis and expression of a Borrelia burgdorferi gene
RT   encoding a 22 kDa protein (pC) in Escherichia coli.";
RL   Mol. Microbiol. 6:503-509(1992).
RN   [2]
RP   SEQUENCE OF 1-205 FROM N.A.
RC   STRAIN=DK26;
RX   MEDLINE=94075528; PubMed=8253951;
RA   Theisen M., Frederiksen B., Lebech A.M., Vuust J., Hansen K.;
RT   "Polymorphism in ospC gene of Borrelia burgdorferi and
RT   immunoreactivity of OspC protein: implications for taxonomy and for
RT   use of OspC protein as a diagnostic antigen.";
RL   J. Clin. Microbiol. 31:2570-2576(1993).
CC   -!- FUNCTION: NOT KNOWN; MAJOR IMMUNODOMINANT PROTEIN.
CC   -!- SUBCELLULAR LOCATION: Attached to the outer membrane by a lipid
CC       anchor.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X62162; CAA44093.1; -.
DR   EMBL; X73624; CAA52003.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
KW   Outer membrane; Lipoprotein; Signal; Plasmid; Antigen.
FT   SIGNAL        1     18       BY SIMILARITY.
FT   CHAIN        19    212       OUTER SURFACE PROTEIN C.
FT   LIPID        19     19       N-ACYL DIGLYCERIDE (BY SIMILARITY).
SQ   SEQUENCE   212 AA;  22499 MW;  C206C231FBF2E7D4 CRC64;


  Query Match             100.0%;  Score 52;  DB 1;  Length 212;
  Best Local Similarity   100.0%;  Pred. No. 0.028;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 PVVAESPKKP 10
          ||||||||||
Db    203 PVVAESPKKP 212


RESULT   3
CLP1_RHIME
ID   CLP1_RHIME     STANDARD;      PRT;   205 AA.
AC   P58277;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   ATP-dependent Clp protease proteolytic subunit 1 (EC 3.4.21.92)
DE   (Endopeptidase Clp 1).
GN   CLPP1 OR R03224 OR SMC03841.
OS   Rhizobium meliloti (Sinorhizobium meliloti).
OC   Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC   Rhizobiaceae; Sinorhizobium.
OX   NCBI_TaxID=382;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=1021;
RX   MEDLINE=21396507; PubMed=11481430;
RA   Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,
RA   Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
RA   Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,
RA   Pohl T., Portetelle D., Puehler A., Purnelle B., Ramsperger U.,
RA   Renard C., Thebault P., Vandenbol M., Weidner S., Galibert F.;
RT   "Analysis of the chromosome sequence of the legume symbiont
RT   Sinorhizobium meliloti strain 1021.";
RL   Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
CC   -!- FUNCTION: Cleaves peptides in various proteins in a process that
CC       requires ATP hydrolysis. Has a chymotrypsin-like activity. Plays a
CC       major role in the degradation of misfolded proteins (By
CC       similarity).
CC   -!- CATALYTIC ACTIVITY: Hydrolysis of proteins to small peptides in
CC       the presence of ATP and magnesium. Alpha-casein is the usual test
CC       substrate. In the absence of ATP, only oligopeptides shorter than
CC       five residues are cleaved (such as succinyl-Leu-Tyr-|-NHMEC; and
CC       Leu-Tyr-Leu-|-Tyr-Trp, in which the cleavage of the -Tyr-|-Leu-
CC       and -Tyr-|-Trp- bond also occurs).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S14; ALSO KNOWN AS CLPP
CC       FAMILY.
```

```
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AL591793; CAC47803.1; -.
DR   InterPro; IPR001907; CLP_protease.
DR   Pfam; PF00574; CLP_protease; 1.
DR   PRINTS; PR00127; CLPPROTEASEP.
DR   PROSITE; PS00382; CLP_PROTEASE_HIS; 1.
DR   PROSITE; PS00381; CLP_PROTEASE_SER; FALSE_NEG.
KW   Hydrolase; Serine protease; Complete proteome.
FT   ACT_SITE    102    102       BY SIMILARITY.
FT   ACT_SITE    127    127       BY SIMILARITY.
SQ   SEQUENCE   205 AA;  23179 MW;  7E2D00CAF1E75B79 CRC64;


  Query Match             71.2%;  Score 37;  DB 1;  Length 205;
  Best Local Similarity   100.0%;  Pred. No. 12;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       4 AESPKKP 10
           |||||||
Db      57 AESPKKP 63


RESULT   4
RP54_AZOVI
ID   RP54_AZOVI     STANDARD;      PRT;   502 AA.
AC   P08623;
DT   01-AUG-1988 (Rel. 08, Created)
DT   01-AUG-1988 (Rel. 08, Last sequence update)
DT   01-FEB-1994 (Rel. 28, Last annotation update)
DE   RNA polymerase sigma-54 factor.
GN   RPON OR NTRA.
OS   Azotobacter vinelandii.
OC   Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC   Azotobacter.
OX   NCBI_TaxID=354;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=UW;
RX   MEDLINE=88142550; PubMed=3481423;
RA   Merrick M.J., Gibbins J., Toukdarian A.;
RT   "The nucleotide sequence of the sigma factor gene ntrA (rpoN) of
RT   Azotobacter vinelandii: analysis of conserved sequences in NtrA
RT   proteins.";
RL   Mol. Gen. Genet. 210:323-330(1987).
CC   -!- FUNCTION: THE SIGMA FACTOR IS AN INITIATION FACTOR THAT PROMOTES
CC       ATTACHMENT OF THE RNA POLYMERASE TO SPECIFIC INITIATION SITES AND
CC       THEN IS RELEASED.
CC   -!- FUNCTION: THIS SIGMA FACTOR IS RESPONSIBLE FOR THE EXPRESSION OF
CC       THE NITROGEN FIXATION PROTEINS. THE OPEN COMPLEX (SIGMA-54 AND
CC       CORE RNA POLYMERASE) SERVES AS THE RECEPTOR FOR RECEIPT OF THE
CC       MELTING SIGNAL FROM THE REMOTELY BOUND ACTIVATOR PROTEIN NIFA FOR
CC       THE EXPRESSION OF THE NITROGEN FIXATION PROTEINS.
CC   -!- SIMILARITY: BELONGS TO THE SIGMA-54 FACTOR FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X05888; CAA29314.1; -.
DR   PIR; S00720; S00720.
DR   InterPro; IPR000394; Sigma54_factor.
DR   Pfam; PF00309; Sigma54_factors; 1.
DR   PRINTS; PR00045; SIGMA54FCT.
DR   PROSITE; PS00717; SIGMA54_1; 1.
DR   PROSITE; PS00718; SIGMA54_2; 1.
DR   PROSITE; PS50044; SIGMA54_3; 1.
KW   Transcription regulation; Sigma factor; DNA-directed RNA polymerase;
KW   DNA-binding; Nitrogen fixation.
FT   DOMAIN       11     38       GLN-RICH.
FT   DOMAIN       19     40       LEUCINE-ZIPPER (POTENTIAL).
FT   DOMAIN      186    207       LEUCINE-ZIPPER (POTENTIAL).
FT   DNA_BIND    391    410       H-T-H MOTIF (POTENTIAL).
FT   SITE        479    487       RPON BOX.
SQ   SEQUENCE   502 AA;  56917 MW;  699A405BE87D4F10 CRC64;


  Query Match             71.2%;  Score 37;  DB 1;  Length 502;
  Best Local Similarity   66.7%;  Pred. No. 29;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 VVAESPKKP 10
           : ||:||||
Db     453 IAAENPKKP 461


RESULT   5
CPN_DROME
ID   CPN_DROME      STANDARD;      PRT;   865 AA.
AC   Q02910;
DT   01-OCT-1993 (Rel. 27, Created)
DT   01-OCT-1993 (Rel. 27, Last sequence update)
DT   01-FEB-1994 (Rel. 28, Last annotation update)
DE   Calphotin.
GN   CPN OR CAP.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CANTON-S;
RX   MEDLINE=93165729; PubMed=8094559;
RA   Martin J.H., Benzer S., Rudnicka M., Miller C.A.;
RT   "Calphotin: a Drosophila photoreceptor cell calcium-binding protein.";
RL   Proc. Natl. Acad. Sci. U.S.A. 90:1531-1535(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CANTON-S;
RX   MEDLINE=93165730; PubMed=8434015;
RA   Ballinger D.G., Xue N., Harshman K.D.;
RT   "A Drosophila photoreceptor cell-specific protein, calphotin, binds
RT   calcium and contains a leucine zipper.";
RL   Proc. Natl. Acad. Sci. U.S.A. 90:1536-1540(1993).
CC   -!- FUNCTION: MIGHT FUNCTION AS A CALCIUM-SEQUESTERING "SPONGE" TO
CC       REGULATE THE AMOUNT OF FREE CYTOPLASMIC CALCIUM. IT BINDS 0.3 MOL
CC       OF CA+2 PER MOL OF PROTEIN.
CC   -!- SUBUNIT: HOMODIMER (PROBABLE).
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC; HYPODENSE COMPARTMENT.
CC   -!- TISSUE SPECIFICITY: SOMA AND AXONS OF PHOTORECEPTOR CELLS OF
CC       COMPOUND EYES AND OCELLI.
CC   -!- DEVELOPMENTAL STAGE: EXPRESSED EARLY IN PHOTORECEPTOR CELL
CC       DEVELOPMENT.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; L02111; AAA28405.1; -.
DR   EMBL; L05080; AAA28420.1; -.
```

```
DR   PIR; A47282; A47282.
DR   FlyBase; FBgn0010218; Cpn.
KW   Calcium-binding.
FT   CONFLICT     36     36       A -> AVAPAVVA (IN REF. 2).
FT   CONFLICT     43     43       I -> T (IN REF. 2).
FT   CONFLICT     64     64       I -> V (IN REF. 2).
FT   CONFLICT     76     76       T -> A (IN REF. 2).
FT   CONFLICT    100    100       P -> PP (IN REF. 2).
FT   CONFLICT    126    127       VQ -> AP (IN REF. 2).
FT   CONFLICT    154    154       I -> V (IN REF. 2).
FT   CONFLICT    160    160       S -> T (IN REF. 2).
FT   CONFLICT    534    534       A -> E (IN REF. 2).
FT   CONFLICT    699    699       I -> T (IN REF. 2).
FT   CONFLICT    703    703       V -> L (IN REF. 2).
FT   CONFLICT    721    721       D -> E (IN REF. 2).
SQ   SEQUENCE   865 AA;  84781 MW;  2110417E0B0E7CFE CRC64;


  Query Match             71.2%;  Score 37;  DB 1;  Length 865;
  Best Local Similarity   70.0%;  Pred. No. 49;
  Matches    7;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 PVVAESPKKP 10
           |||||:|  |
Db     200 PVVAETPAPP 209


RESULT   6
YRB1_SYNP6
ID   YRB1_SYNP6     STANDARD;      PRT;   122 AA.
AC   P23655;
DT   01-NOV-1991 (Rel. 20, Created)
DT   01-NOV-1991 (Rel. 20, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Hypothetical 12.9 kDa protein in RuBisCO 5'region (ORF1).
OS   Synechococcus sp. (strain PCC 6301) (Anacystis nidulans).
OC   Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
OX   NCBI_TaxID=1139;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Shinozaki K., Sugiura M.;
RT   "Genes for the large and small subunits of ribulose-1,5-bisphosphate
RT   carboxylase/oxygenase constitute a single operon in a cyanobacterium
RT   Anacystis nidulans 6301.";
RL   Mol. Gen. Genet. 200:27-32(1985).
CC   -!- FUNCTION: MAY BE INVOLVED IN THE FORMATION OF THE CARBOXYSOME, A
CC       POLYHEDRAL INCLUSION WHERE RUBISCO IS SEQUESTERED.
CC   -!- SIMILARITY: BELONGS TO THE CCHA/CCMK/CSOS1/PDUA FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X03220; CAA26970.1; -.
DR   PIR; S07310; S07310.
DR   InterPro; IPR000249; Bact_microcomp.
DR   Pfam; PF00936; Bact_microcomp; 1.
DR   ProDom; PD003442; Bact_microcomp; 1.
DR   PROSITE; PS01139; BACT_MICROCOMP; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   122 AA;  12855 MW;  F54244DE24531504 CRC64;


  Query Match             69.2%;  Score 36;  DB 1;  Length 122;
  Best Local Similarity   60.0%;  Pred. No. 11;
  Matches    6;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 PVVAESPKKP 10
           |:   |||||
Db     100 PIAGGSPKKP 109


RESULT   7
YJO7_YEAST
ID   YJO7_YEAST     STANDARD;      PRT;   382 AA.
AC   P47007;
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   01-FEB-1996 (Rel. 33, Last annotation update)
DE   Hypothetical 44.9 kDa protein in INO1-IDS2 intergenic region.
GN   YJL147C OR J0639.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288C / FY1679;
RX   MEDLINE=96408771; PubMed=8813765;
RA   Katsoulou C., Tzermia M., Tavernarakis N., Alexandraki D.;
RT   "Sequence analysis of a 40.7 kb segment from the left arm of yeast
RT   chromosome X reveals 14 known genes and 13 new open reading frames
RT   including homologues of genes clustered on the right arm of
RT   chromosome XI.";
RL   Yeast 12:787-797(1996).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z49422; CAA89442.1; -.
DR   EMBL; X87371; CAA60808.1; -.
DR   SGD; S0003683; YJL147C.
KW   Hypothetical protein.
SQ   SEQUENCE   382 AA;  44862 MW;  DDAAF588AF9A3234 CRC64;


  Query Match             69.2%;  Score 36;  DB 1;  Length 382;
  Best Local Similarity   66.7%;  Pred. No. 34;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 VVAESPKKP 10
           :: ||||||
Db     313 IMLESPKKP 321


RESULT   8
SDC3_CAEEL
ID   SDC3_CAEEL     STANDARD;      PRT;  2150 AA.
AC   P34706;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Zinc finger protein sdc-3.
GN   SDC-3.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RX   MEDLINE=93161411; PubMed=8431944;
RA   Klein R.D., Meyer B.J.;
RT   "Independent domains of the Sdc-3 protein control sex determination
RT   and dosage compensation in C. elegans.";
RL   Cell 72:349-364(1993).
CC   -!- FUNCTION: CONTROLS BOTH SEX DETERMINATION AND X CHROMOSOME DOSAGE
```

```
CC       COMPENSATION. THESE TWO FUNCTIONS ACT INDEPENDENTLY.
CC    -!- SUBCELLULAR LOCATION: Nuclear.
CC    -!- DEVELOPMENTAL STAGE: EXPRESSED IN EMBRYONIC AND EARLY LARVAL
CC        STAGES.
CC    ---------------------------------------------------------------------
CC    This SWISS-PROT entry is copyright. It is produced through a collaboration
CC    between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC    the European Bioinformatics Institute. There are no restrictions on its
CC    use by non-profit institutions as long as its content is in no way
CC    modified and this statement is not removed. Usage by and for commercial
CC    entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC    or send an email to license@isb-sib.ch).
CC    ---------------------------------------------------------------------
DR    EMBL; M85149; AAA28144.1; -.
DR    PIR; S27802; S27802.
DR    InterPro; IPR000822; Znf-C2H2.
DR    PROSITE; PS00028; ZINC_FINGER_C2H2_1; 1.
DR    PROSITE; PS50157; ZINC_FINGER_C2H2_2; FALSE_NEG.
KW    Developmental protein; Zinc-finger; Metal-binding; DNA-binding;
KW    Nuclear protein; Repeat.
FT    DOMAIN      443    987       DOSAGE COMPENSATION DOMAIN 1.
FT    DOMAIN     1508   1516       SEX DETERMINATION DOMAIN.
FT    DOMAIN     2080   2105       DOSAGE COMPENSATION DOMAIN 2.
FT    ZN_FING    2078   2105       C2H2-TYPE.
FT    ZN_FING    2117   2141       C2H2-TYPE.
SQ    SEQUENCE   2150 AA;  249954 MW;  7430D77AC784EA46 CRC64;


  Query Match             69.2%;  Score 36;  DB 1;  Length 2150;
  Best Local Similarity   77.8%;  Pred. No. 1.8e+02;
  Matches    7;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 PVVAESPKK 9
           ||  |||||
Db    1313 PVAEESPKK 1321


RESULT   9
HMX1_BOVIN
ID    HMX1_BOVIN     STANDARD;      PRT;   297 AA.
AC    O02786;
DT    15-JUL-1998 (Rel. 36, Created)
DT    15-JUL-1998 (Rel. 36, Last sequence update)
DT    15-JUL-1998 (Rel. 36, Last annotation update)
DE    Homeobox protein MSX-1.
GN    MSX1.
OS    Bos taurus (Bovine).
OC    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC    Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC    Bovidae; Bovinae; Bos.
OX    NCBI_TaxID=9913;
RN    [1]
RP    SEQUENCE FROM N.A.
RC    TISSUE=Tooth;
RX    MEDLINE=95352824; PubMed=7626784;
RA    Iimura T., Oida S., Takeda K., Maruoka Y., Shimokawa H., Ibaraki K.,
RA    Sasaki S.;
RT    "Molecular cloning and sequence of bovine Msx-1 homeobox-containing
RT    gene cDNA from a bovine odontoblast library.";
RL    DNA Seq. 5:233-237(1995).
CC    -!- FUNCTION: PROBABLE MORPHOGENETIC ROLE. MAY PLAY A ROLE IN LIMB-
CC        PATTERN FORMATION (BY SIMILARITY).
CC    -!- SUBCELLULAR LOCATION: Nuclear.
CC    -!- SIMILARITY: BELONGS TO THE MSH FAMILY OF HOMEOBOX PROTEINS.
CC    ---------------------------------------------------------------------
CC    This SWISS-PROT entry is copyright. It is produced through a collaboration
CC    between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC    the European Bioinformatics Institute. There are no restrictions on its
CC    use by non-profit institutions as long as its content is in no way
CC    modified and this statement is not removed. Usage by and for commercial
CC    entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC    or send an email to license@isb-sib.ch).
CC    ---------------------------------------------------------------------
DR    EMBL; D30750; BAA20367.1; -.
DR    HSSP; P22808; 1NK3.
DR    InterPro; IPR001356; Homeobox.
DR    Pfam; PF00046; homeobox; 1.
DR    PRINTS; PR00024; HOMEOBOX.
DR    SMART; SM00389; HOX; 1.
DR    PROSITE; PS00027; HOMEOBOX_1; 1.
DR    PROSITE; PS50071; HOMEOBOX_2; 1.
KW    Homeobox; DNA-binding; Developmental protein; Nuclear protein.
FT    DNA_BIND    166    225       HOMEOBOX.
SQ    SEQUENCE   297 AA;  31111 MW;  4D0C987EE1F6BFBE CRC64;


  Query Match             67.3%;  Score 35;  DB 1;  Length 297;
  Best Local Similarity   77.8%;  Pred. No. 40;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 VVAESPKKP 10
           | ||||:||
Db     126 VKAESPEKP 134


RESULT  10
HMX1_HUMAN
ID    HMX1_HUMAN     STANDARD;      PRT;   297 AA.
AC    P28360; Q96NY4;
DT    01-DEC-1992 (Rel. 24, Created)
DT    01-DEC-1992 (Rel. 24, Last sequence update)
DT    01-MAR-2002 (Rel. 41, Last annotation update)
DE    Homeobox protein MSX-1 (Hox-7).
GN    MSX1 OR HOX7.
OS    Homo sapiens (Human).
OC    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX    NCBI_TaxID=9606;
RN    [1]
RP    SEQUENCE FROM N.A.
RX    MEDLINE=93250782; PubMed=1284527;
RA    Padanilam B.J., Stadler S.H., Mills K.A., McLeod L.B., Solursh M.,
RA    Lee B.M., Ramirez F., Buetow K.H., Murray J.C.;
RT    "Characterization of the human HOX 7 cDNA and identification of
RT    polymorphic markers.";
RL    Hum. Mol. Genet. 1:407-410(1992).
RN    [2]
RP    SEQUENCE FROM N.A.
RX    MEDLINE=92128949; PubMed=1685479;
RA    Hewitt J.E., Clarke L.E., Iven A., Williamson R.;
RT    "Structure and sequence of the human homeobox gene HOX7.";
RL    Genomics 11:670-678(1991).
RN    [3]
RP    SEQUENCE FROM N.A.
RA    Jezewski P.A., Vieira A.;
RT    "Diagnostic resequencing demonstrates a role for MSX1 in nonsyndromic
RT    cleft lip and palate.";
RL    Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
RN    [4]
RP    VARIANT FTA PRO-196.
RX    MEDLINE=96331281; PubMed=8696335;
RA    Vastardis H., Karimbux N., Guthua S.W., Seidman J.G., Seidman C.E.;
RT    "A human MSX1 homeodomain missense mutation causes selective tooth
RT    agenesis.";
RL    Nat. Genet. 13:417-421(1996).
CC    -!- FUNCTION: PROBABLE MORPHOGENETIC ROLE. MAY PLAY A ROLE IN LIMB-
CC        PATTERN FORMATION.
CC    -!- SUBCELLULAR LOCATION: Nuclear.
CC    -!- DISEASE: IMPLICATED IN WOLF-HIRSCHORN SYNDROME (WHS), WHICH IS
CC        CHARACTERIZED BY PROFOUND MENTAL RETARDATION, HEART DEFECTS, AND
CC        FACIAL CLEFTING.
CC    -!- DISEASE: DEFECTS IN MSX1 ARE IMPLICATED IN FAMILIAL TOOTH AGENESIS
CC        (FTA). AGENESIS OF ONE OR MORE TEETH CONSTITUTES ONE OF THE MOST
CC        COMMON DEVELOPMENTAL ANOMALIES IN MAN. REPORTED INCIDENCES VARY
CC        FROM 1.6% TO 9.6%, EXCLUDING THIRD MOLAR (WISDOM TOOTH) AGENESIS,
CC        WHICH OCCURS IN 20% OF THE POPULATION.
```

```
CC   -!- SIMILARITY: BELONGS TO THE MSH FAMILY OF HOMEOBOX PROTEINS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M97676; AAA52683.1; -.
DR   EMBL; M76732; AAA58665.1; -.
DR   EMBL; M76731; AAA58665.1; JOINED.
DR   EMBL; AF426432; AAL17870.1; -.
DR   PIR; A40560; A40560.
DR   HSSP; P22808; 1NK3.
DR   TRANSFAC; T02071; -.
DR   MIM; 142983; -.
DR   MIM; 106600; -.
DR   InterPro; IPR001827; Antennapedia.
DR   InterPro; IPR001356; Homeobox.
DR   Pfam; PF00046; homeobox; 1.
DR   PRINTS; PR00025; ANTENNAPEDIA.
DR   PRINTS; PR00024; HOMEOBOX.
DR   SMART; SM00389; HOX; 1.
DR   PROSITE; PS00027; HOMEOBOX_1; 1.
DR   PROSITE; PS50071; HOMEOBOX_2; 1.
KW   Homeobox; DNA-binding; Developmental protein; Nuclear protein;
KW   Disease mutation.
FT   DNA_BIND     166    225       HOMEOBOX.
FT   VARIANT      196    196       R -> P (IN FTA).
FT                                 /FTId=VAR_003754.
FT   CONFLICT      39     39       A -> T (IN REF. 2).
FT   CONFLICT      91     92       GV -> AS (IN REF. 2).
FT   CONFLICT      93     93       R -> P (IN REF. 1 AND 3).
SQ   SEQUENCE   297 AA;  30998 MW;  A19B4F0E39567B41 CRC64;


  Query Match             67.3%;  Score 35;  DB 1;  Length 297;
  Best Local Similarity   77.8%;  Pred. No. 40;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 VVAESPKKP 10
           | ||||:||
Db     126 VKAESPEKP 134


RESULT  11
ODC_HUMAN
ID   ODC_HUMAN      STANDARD;      PRT;   299 AA.
AC   Q9BQT8;
DT   01-MAR-2002 (Rel. 41, Created)
DT   01-MAR-2002 (Rel. 41, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Mitochondrial 2-oxodicarboxylate carrier.
GN   SLC25A21 OR ODC.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A., CHARACTERIZATION, AND TISSUE SPECIFICITY.
RC   TISSUE=Liver;
RX   MEDLINE=21269385; PubMed=11083877;
RA   Fiermonte G., Dolce V., Palmieri L., Ventura M., Runswick M.J.,
RA   Palmieri F., Walker J.E.;
RT   "Identification of the human mitochondrial oxodicarboxylate carrier:
RT   Bacterial expression, reconstitution, functional characterization,
RT   tissue distribution and chromosomal location.";
RL   J. Biol. Chem. 276:8225-8230(2001).
CC   -!- FUNCTION: Transports C5-C7 oxodicarboxylates across the inner
CC       membranes of mitochondria. Can transport 2-oxoadipate, 2-
CC       oxoglutarate, adipate, glutarate, and to a lesser extent,
CC       pimelate, 2-oxopimelate, 2-aminoadipate, oxaloacetate, and
CC       citrate.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC       inner membrane.
CC   -!- DOMAIN: COMPOSED OF THREE HOMOLOGOUS DOMAINS.
CC   -!- TISSUE SPECIFICITY: Ubiquitous.
CC   -!- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AJ278148; CAC27562.1; -.
DR   InterPro; IPR001993; Mitoch_carrier.
DR   InterPro; IPR002067; Mit_carrier.
DR   Pfam; PF00153; mito_carr; 3.
DR   PRINTS; PR00926; MITOCARRIER.
DR   PROSITE; PS00215; MITOCH_CARRIER; 2.
KW   Mitochondrion; Inner membrane; Repeat; Transmembrane; Transport.
FT   TRANSMEM     17     37       POTENTIAL.
FT   TRANSMEM     62     82       POTENTIAL.
FT   TRANSMEM    100    120       POTENTIAL.
FT   TRANSMEM    211    231       POTENTIAL.
SQ   SEQUENCE   299 AA;  33303 MW;  69A259400328AE19 CRC64;


  Query Match             67.3%;  Score 35;  DB 1;  Length 299;
  Best Local Similarity   55.6%;  Pred. No. 40;
  Matches    5;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 PVVAESPKK 9
             |::||:||:
Db      79 PILAETPKR 87


RESULT  12
HRP3_SCHPO
ID   HRP3_SCHPO      STANDARD;      PRT;  1388 AA.
AC   O14139;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Chromodomain helicase hrp3.
GN   HRP3 OR SPAC3G6.01.
OS   Schizosaccharomyces pombe (Fission yeast).
OC   Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC   Schizosaccharomycetales; Schizosaccharomycetaceae;
OC   Schizosaccharomyces.
OX   NCBI_TaxID=4896;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=972;
RA   Gentles S., Churcher C.M., Barrell B.G., Rajandream M.A., Wood V.;
RL   Submitted (SEP-1997) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   GENE NAME.
RA   Bjerling P., Ekwall K.;
RL   Submitted (MAR-2001) to the SWISS-PROT data bank.
CC   -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC   -!- SIMILARITY: BELONGS TO THE SNF2/RAD54 HELICASE FAMILY.
CC   -!- SIMILARITY: CONTAINS 2 'CHROMO' DOMAINS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
```

```
CC   ---------------------------------------------------------------------
DR   EMBL; Z99167; CAB16277.1; -.
DR   InterPro; IPR000953; Chromo.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR001650; Helicase_C.
DR   InterPro; IPR000330; SNF2_N.
DR   Pfam; PF00385; chromo; 2.
DR   Pfam; PF00271; helicase_C; 1.
DR   Pfam; PF00176; SNF2_N; 1.
DR   SMART; SM00298; CHROMO; 2.
DR   SMART; SM00487; DEXDc; 1.
DR   SMART; SM00490; HELICc; 1.
DR   PROSITE; PS00598; CHROMO_1; 1.
DR   PROSITE; PS50013; CHROMO_2; 2.
KW   ATP-binding; Helicase; DNA-binding; Nuclear protein; Repeat.
FT   DOMAIN      169    175       POLY-GLU.
FT   DOMAIN      191    260       CHROMO 1.
FT   DOMAIN      288    349       CHROMO 2.
FT   NP_BIND     400    407       ATP (POTENTIAL).
FT   SITE        508    511       DEAH BOX.
FT   DOMAIN      947    950       POLY-GLU.
SQ   SEQUENCE   1388 AA;  159377 MW;  F7B431084BD29F8D CRC64;


  Query Match             67.3%;  Score 35;  DB 1;  Length 1388;
  Best Local Similarity   60.0%;  Pred. No. 1.8e+02;
  Matches    6;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 PVVAESPKKP 10
           | ::|| |||
Db    1251 PAISESRKKP 1260


RESULT  13
VE7_PAPVD
ID   VE7_PAPVD      STANDARD;      PRT;   102 AA.
AC   P03131;
DT   21-JUL-1986 (Rel. 01, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   E7 protein.
GN   E7.
OS   Deer papillomavirus.
OC   Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC   Papillomavirus.
OX   NCBI_TaxID=10564;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=85293253; PubMed=2993669;
RA   Groff D.E., Lancaster W.D.;
RT   "Molecular cloning and nucleotide sequence of deer papillomavirus.";
RL   J. Virol. 56:85-91(1985).
CC   -!- FUNCTION: E7 PROTEIN HAS BOTH TRANSFORMING AND TRANS-ACTIVATING
CC       ACTIVITIES.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M11910; AAA66842.1; -.
DR   PIR; A03693; W7WLDP.
DR   InterPro; IPR000148; Papvi_E7.
DR   Pfam; PF00527; E7; 1.
KW   Early protein; Transcription regulation; Oncogene;
KW   DNA-binding; Trans-acting factor.
FT   DOMAIN       61     64       C-XX-C MOTIF 1.
FT   DOMAIN       94     97       C-XX-C MOTIF 2.
SQ   SEQUENCE   102 AA;  11166 MW;  CEA837C8955FC003 CRC64;


  Query Match             65.4%;  Score 34;  DB 1;  Length 102;
  Best Local Similarity   60.0%;  Pred. No. 22;
  Matches    6;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 PVVAESPKKP 10
           ||| : | ||
Db      37 PVVVDKPGKP 46


RESULT  14
PAB_PEPMA
ID   PAB_PEPMA      STANDARD;      PRT;   387 AA.
AC   Q51911;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   Peptostreptococcal albumin-binding protein precursor.
GN   PAB.
OS   Peptostreptococcus magnus.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC   Finegoldia.
OX   NCBI_TaxID=1260;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ALB8;
RX   MEDLINE=94216330; PubMed=8163519;
RA   de Chateau M., Bjoerck L.;
RT   "Protein PAB, a mosaic albumin-binding bacterial protein representing
RT   the first contemporary example of module shuffling.";
RL   J. Biol. Chem. 269:12147-12151(1994).
RN   [2]
RP   STRUCTURE BY NMR OF 213-265, AND REVISION TO 244.
RC   STRAIN=ALB8;
RX   MEDLINE=97240805; PubMed=9086265;
RA   Johansson M.U., de Chateau M., Wikstroem M., Forsen S., Drakenberg T.,
RA   Bjoerck L.;
RT   "Solution structure of the albumin-binding GA module: a versatile
RT   bacterial protein domain.";
RL   J. Mol. Biol. 266:859-865(1997).
CC   -!- FUNCTION: BINDS SERUM ALBUMIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X77864; CAA54857.1; -.
DR   PDB; 1GAB; 07-JUL-97.
DR   PDB; 1PRB; 23-JUL-97.
DR   InterPro; IPR002988; GA.
DR   InterPro; IPR001899; Gram_pos_anchor.
DR   Pfam; PF01468; GA; 2.
DR   Pfam; PF00746; Gram_pos_anchor; 1.
KW   Signal; 3D-structure.
FT   SIGNAL        1     26
FT   CHAIN        27    387       PEPTOSTREPTOCOCCAL ALBUMIN-BINDING
FT                                 PROTEIN.
FT   DOMAIN      213    265       GA MODULE.
SQ   SEQUENCE   387 AA;  43057 MW;  3D5135C4C3A3D8F2 CRC64;


  Query Match             65.4%;  Score 34;  DB 1;  Length 387;
  Best Local Similarity   85.7%;  Pred. No. 78;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       4 AESPKKP 10
           ||:||||
Db     194 AETPKKP 200
```

```
RESULT  15
SYTB_MOUSE
ID   SYTB_MOUSE     STANDARD;      PRT;   430 AA.
AC   Q9R0N3;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Synaptotagmin XI (SytXI).
GN   SYT11.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ICR; TISSUE=Cerebellum;
RX   MEDLINE=20002669; PubMed=10531343;
RA   Fukuda M., Kanno E., Mikoshiba K.;
RT   "Conserved N-terminal cysteine motif is essential for homo- and
RT   heterodimer formation of synaptotagmins III, V, VI, and X.";
RL   J. Biol. Chem. 274:31421-31427(1999).
CC   -!- FUNCTION: MAY BE INVOLVED IN CA2+-DEPENDENT EXOCYTOSIS OF
CC       SECRETORY VESICLES THROUGH CA2+ AND PHOSPHOLIPID BINDING TO THE C2
CC       DOMAIN OR MAY SERVE AS CA2+ SENSORS IN THE PROCESS OF VESICULAR
CC       TRAFFICKING AND EXOCYTOSIS (By similarity).
CC   -!- SUBUNIT: Homodimer. Can also forms heterodimer (By similarity).
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. SYNAPTIC
CC       VESICLES (By similarity).
CC   -!- SIMILARITY: CONTAINS 2 C2 DOMAINS.
CC   -!- SIMILARITY: BELONGS TO THE SYNAPTOTAGMIN FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AB026808; BAA85780.1; -.
DR   HSSP; P21707; 1RSY.
DR   MGD; MGI:1859547; Syt11.
DR   InterPro; IPR000008; C2.
DR   InterPro; IPR002149; LRI.
DR   InterPro; IPR001565; Synaptotagmin.
DR   Pfam; PF00168; C2; 2.
DR   PRINTS; PR00360; C2DOMAIN.
DR   PRINTS; PR00399; SYNAPTOTAGMN.
DR   SMART; SM00239; C2; 2.
DR   PROSITE; PS00499; C2_DOMAIN_1; FALSE_NEG.
DR   PROSITE; PS50004; C2_DOMAIN_2; 2.
KW   Transmembrane; Repeat; Synapse.
FT   DOMAIN        1     15       VESICULAR (POTENTIAL).
FT   TRANSMEM     16     36       POTENTIAL.
FT   DOMAIN       37    430       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      173    261       C2 DOMAIN 1.
FT   DOMAIN      303    396       C2 DOMAIN 2.
SQ   SEQUENCE   430 AA;  48359 MW;  25E7CDFC4B4BE036 CRC64;


  Query Match             65.4%;  Score 34;  DB 1;  Length 430;
  Best Local Similarity   75.0%;  Pred. No. 86;
  Matches    6;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       3 VAESPKKP 10
           | |||:||
Db     413 VCESPRKP 420


Search completed: October 12, 2002, 20:48:01
Job time: 320 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: October 12, 2002, 17:27:46 ; Search time 87.59 Seconds
(without alignments)
19.751 Million cell updates/sec

Title: US-09-408-578A-1
Perfect score: 52
Sequence: 1 PVVAESPKKP 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_19:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	52	100.0	190	2	P94244	P94244 borrelia bu
2	52	100.0	190	2	P70819	P70819 borrelia bu
3	52	100.0	191	2	P94245	P94245 borrelia bu
4	52	100.0	191	2	Q9S3P0	Q9s3p0 borrelia bu
5	52	100.0	191	2	P70818	P70818 borrelia bu
6	52	100.0	192	2	Q9S3P3	Q9s3p3 borrelia bu
7	52	100.0	192	2	Q9S3P2	Q9s3p2 borrelia bu
8	52	100.0	193	2	P94226	P94226 borrelia bu
9	52	100.0	193	2	P94233	P94233 borrelia bu
10	52	100.0	193	2	P94237	P94237 borrelia bu
11	52	100.0	194	2	P94247	P94247 borrelia bu
12	52	100.0	194	2	Q9S3P4	Q9s3p4 borrelia bu
13	52	100.0	194	2	P94229	P94229 borrelia bu
14	52	100.0	194	2	Q45030	Q45030 borrelia bu
15	52	100.0	207	2	Q45187	Q45187 borrelia ga
16	52	100.0	207	2	Q45177	Q45177 borrelia ga
17	52	100.0	207	2	Q49581	Q49581 borrelia ga
18	52	100.0	207	2	Q45175	Q45175 borrelia ga
19	52	100.0	207	2	Q07336	Q07336 borrelia bu
20	52	100.0	209	2	Q44883	Q44883 borrelia bu
21	52	100.0	209	2	Q45179	Q45179 borrelia ga
22	52	100.0	209	2	Q9KIK3	Q9kik3 borrelia ga
23	52	100.0	209	2	Q44671	Q44671 borrelia af
24	52	100.0	209	2	Q49579	Q49579 borrelia ga
25	52	100.0	209	2	Q49583	Q49583 borrelia ga
26	52	100.0	209	2	Q49584	Q49584 borrelia ga
27	52	100.0	209	2	P70891	P70891 borrelia ga
28	52	100.0	210	2	Q57279	Q57279 borrelia ga
29	52	100.0	210	2	Q57359	Q57359 borrelia bu
30	52	100.0	210	2	Q9KIM6	Q9kim6 borrelia ga
31	52	100.0	210	2	Q45176	Q45176 borrelia ga
32	52	100.0	210	2	Q45178	Q45178 borrelia ga
33	52	100.0	210	2	Q49582	Q49582 borrelia ga
34	52	100.0	210	2	P70893	P70893 borrelia ga
35	52	100.0	210	2	Q44719	Q44719 borrelia bu
36	52	100.0	210	2	Q44978	Q44978 borrelia bu
37	52	100.0	211	2	Q57262	Q57262 borrelia bu
38	52	100.0	211	2	Q49576	Q49576 borrelia af
39	52	100.0	211	2	Q49577	Q49577 borrelia af
40	52	100.0	211	2	Q44976	Q44976 borrelia bu
41	52	100.0	211	2	Q44720	Q44720 borrelia bu
42	52	100.0	211	2	Q44977	Q44977 borrelia bu
43	52	100.0	211	2	Q9Z6C7	Q9z6c7 borrelia va
44	52	100.0	212	2	Q9KIM5	Q9kim5 borrelia af
45	52	100.0	212	2	Q9KIM4	Q9kim4 borrelia af

ALIGNMENTS

RESULT 1
P94244
ID P94244 PRELIMINARY; PRT; 190 AA.
AC P94244;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE OUTER SURFACE PROTEIN C (FRAGMENT).
GN OSPC.
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VSDA;
RX MEDLINE=96296448; PubMed=8709845;
RA Livey_I., Gibbs C.P., Schuster R., Dorner F.;
RT "Evidence for lateral transfer and recombination in OspC variation in
RT Lyme disease Borrelia.";
RL Mol. Microbiol. 18:257-269(1995).
DR EMBL; L42870; AAB37013.1; -.
DR InterPro; IPR001800; Lipoprotein_6.
DR Pfam; PF01441; Lipoprotein_6; 1.
DR ProDom; PD001149; Lipoprotein_6; 1.
FT NON_TER 1 1
FT NON_TER 190 190
SQ SEQUENCE 190 AA; 19904 MW; 9A0C47B29B73AD8D CRC64;

Query Match 100.0%; Score 52; DB 2; Length 190;
Best Local Similarity 100.0%; Pred. No. 0.076;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PVVAESPKKP 10
||||||||||
Db 181 PVVAESPKKP 190

RESULT 2

```
P70819
ID   P70819      PRELIMINARY;      PRT;   190 AA.
AC   P70819;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   OUTER SURFACE PROTEIN (FRAGMENT).
GN   OSPC.
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=2-1498 SON 188;
RA   Probert W.S., Crawford M.R., Cadiz R.B., LeFebvre R.B.;
RT   "Immuniuzation with OspA, but not ospC, provides protection of mice
RT   challenged with North American isolates of Borrelia burgdorferi.";
RL   Submitted (SEP-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; L81130; AAB06570.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
FT   NON_TER       1      1
SQ   SEQUENCE   190 AA;  19695 MW;  FBB92D4494F2D49A CRC64;


  Query Match             100.0%;  Score 52;  DB 2;  Length 190;
  Best Local Similarity   100.0%;  Pred. No. 0.076;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 PVVAESPKKP 10
          ||||||||||
Db    181 PVVAESPKKP 190


RESULT   3
P94245
ID   P94245      PRELIMINARY;      PRT;   191 AA.
AC   P94245;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C (FRAGMENT).
GN   OSPC.
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=VS461;
RX   MEDLINE=96296448; PubMed=8709845;
RA   Livey I., Gibbs C.P., Schuster R., Dorner F.;
RT   "Evidence for lateral transfer and recombination in OspC variation in
RT   Lyme disease Borrelia.";
RL   Mol. Microbiol. 18:257-269(1995).
DR   EMBL; L42871; AAB37014.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
FT   NON_TER       1      1
FT   NON_TER     191    191
SQ   SEQUENCE   191 AA;  19923 MW;  2C2F3BD40714EAA8 CRC64;


  Query Match             100.0%;  Score 52;  DB 2;  Length 191;
  Best Local Similarity   100.0%;  Pred. No. 0.076;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 PVVAESPKKP 10
          ||||||||||
Db    182 PVVAESPKKP 191


RESULT   4
Q9S3P0
ID   Q9S3P0      PRELIMINARY;      PRT;   191 AA.
AC   Q9S3P0;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C (FRAGMENT).
GN   OSPC.
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=27579;
RX   MEDLINE=96296448; PubMed=8709845;
RA   Livey I., Gibbs C.P., Schuster R., Dorner F.;
RT   "Evidence for lateral transfer and recombination in OspC variation in
RT   Lyme disease Borrelia.";
RL   Mol. Microbiol. 18:257-269(1995).
DR   EMBL; L42896; AAB37004.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
FT   NON_TER       1      1
FT   NON_TER     191    191
SQ   SEQUENCE   191 AA;  19826 MW;  DBA79667814F290A CRC64;


  Query Match             100.0%;  Score 52;  DB 2;  Length 191;
  Best Local Similarity   100.0%;  Pred. No. 0.076;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 PVVAESPKKP 10
          ||||||||||
Db    182 PVVAESPKKP 191


RESULT   5
P70818
ID   P70818      PRELIMINARY;      PRT;   191 AA.
AC   P70818;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   OUTER SURFACE PROTEIN (FRAGMENT).
GN   OSPC.
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=2-1498 CA4;
RA   Probert W.S., Crawford M.R., Cadiz R.B., LeFebvre R.B.;
RT   "Immuniuzation with OspA, but not ospC, provides protection of mice
RT   challenged with North American isolates of Borrelia burgdorferi.";
RL   Submitted (SEP-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; L81131; AAB06569.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
FT   NON_TER       1      1
SQ   SEQUENCE   191 AA;  20126 MW;  D2B9B1C82B4DC3C0 CRC64;


  Query Match             100.0%;  Score 52;  DB 2;  Length 191;
  Best Local Similarity   100.0%;  Pred. No. 0.076;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 PVVAESPKKP 10
          ||||||||||
Db    182 PVVAESPKKP 191
```

```
RESULT   6
Q9S3P3
ID   Q9S3P3      PRELIMINARY;      PRT;   192 AA.
AC   Q9S3P3;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C (FRAGMENT).
GN   OSPC.
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=IP2;
RX   MEDLINE=96296448; PubMed=8709845;
RA   Livey I., Gibbs C.P., Schuster R., Dorner F.;
RT   "Evidence for lateral transfer and recombination in OspC variation in
RT   Lyme disease Borrelia.";
RL   Mol. Microbiol. 18:257-269(1995).
DR   EMBL; L42887; AAB36995.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
FT   NON_TER       1      1
FT   NON_TER     192    192
SQ   SEQUENCE   192 AA;  20287 MW;  11846F7AC84C7E3D CRC64;


  Query Match             100.0%;  Score 52;  DB 2;  Length 192;
  Best Local Similarity   100.0%;  Pred. No. 0.077;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 PVVAESPKKP 10
           ||||||||||
Db     183 PVVAESPKKP 192


RESULT   7
Q9S3P2
ID   Q9S3P2      PRELIMINARY;      PRT;   192 AA.
AC   Q9S3P2;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C (FRAGMENT).
GN   OSPC.
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=297;
RX   MEDLINE=96296448; PubMed=8709845;
RA   Livey I., Gibbs C.P., Schuster R., Dorner F.;
RT   "Evidence for lateral transfer and recombination in OspC variation in
RT   Lyme disease Borrelia.";
RL   Mol. Microbiol. 18:257-269(1995).
DR   EMBL; L42893; AAB37001.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
FT   NON_TER       1      1
FT   NON_TER     192    192
SQ   SEQUENCE   192 AA;  20472 MW;  46AC8F93E4DFED6C CRC64;


  Query Match             100.0%;  Score 52;  DB 2;  Length 192;
  Best Local Similarity   100.0%;  Pred. No. 0.077;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 PVVAESPKKP 10
           ||||||||||
Db     183 PVVAESPKKP 192


RESULT   8
P94226
ID   P94226      PRELIMINARY;      PRT;   193 AA.
AC   P94226;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C (FRAGMENT).
GN   OSPC.
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=J1;
RX   MEDLINE=96296448; PubMed=8709845;
RA   Livey I., Gibbs C.P., Schuster R., Dorner F.;
RT   "Evidence for lateral transfer and recombination in OspC variation in
RT   Lyme disease Borrelia.";
RL   Mol. Microbiol. 18:257-269(1995).
DR   EMBL; L42884; AAB36992.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
FT   NON_TER       1      1
FT   NON_TER     193    193
SQ   SEQUENCE   193 AA;  20125 MW;  DF3F926F8BF07290 CRC64;


  Query Match             100.0%;  Score 52;  DB 2;  Length 193;
  Best Local Similarity   100.0%;  Pred. No. 0.077;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 PVVAESPKKP 10
           ||||||||||
Db     184 PVVAESPKKP 193


RESULT   9
P94233
ID   P94233      PRELIMINARY;      PRT;   193 AA.
AC   P94233;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C (FRAGMENT).
GN   OSPC.
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ACA1;
RX   MEDLINE=96296448; PubMed=8709845;
RA   Livey I., Gibbs C.P., Schuster R., Dorner F.;
RT   "Evidence for lateral transfer and recombination in OspC variation in
RT   Lyme disease Borrelia.";
RL   Mol. Microbiol. 18:257-269(1995).
DR   EMBL; L42892; AAB37000.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
FT   NON_TER       1      1
FT   NON_TER     193    193
SQ   SEQUENCE   193 AA;  20370 MW;  CD52748E2EA2D36F CRC64;


  Query Match             100.0%;  Score 52;  DB 2;  Length 193;
  Best Local Similarity   100.0%;  Pred. No. 0.077;
```

```
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 PVVAESPKKP 10
          ||||||||||
Db    184 PVVAESPKKP 193


RESULT  10
P94237
ID   P94237      PRELIMINARY;      PRT;   193 AA.
AC   P94237;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C (FRAGMENT).
GN   OSPC.
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=25015;
RX   MEDLINE=96296448; PubMed=8709845;
RA   Livey I., Gibbs C.P., Schuster R., Dorner F.;
RT   "Evidence for lateral transfer and recombination in OspC variation in
RT   Lyme disease Borrelia.";
RL   Mol. Microbiol. 18:257-269(1995).
DR   EMBL; L42898; AAB37007.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
FT   NON_TER       1      1
FT   NON_TER     193    193
SQ   SEQUENCE   193 AA;  20677 MW;  C9500D959E13590D CRC64;


  Query Match            100.0%;  Score 52;  DB 2;  Length 193;
  Best Local Similarity  100.0%;  Pred. No. 0.077;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 PVVAESPKKP 10
          ||||||||||
Db    184 PVVAESPKKP 193


RESULT  11
P94247
ID   P94247      PRELIMINARY;      PRT;   194 AA.
AC   P94247;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C (FRAGMENT).
GN   OSPC.
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SIMON;
RX   MEDLINE=96296448; PubMed=8709845;
RA   Livey I., Gibbs C.P., Schuster R., Dorner F.;
RT   "Evidence for lateral transfer and recombination in OspC variation in
RT   Lyme disease Borrelia.";
RL   Mol. Microbiol. 18:257-269(1995).
DR   EMBL; L42873; AAB37016.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
FT   NON_TER       1      1
FT   NON_TER     194    194
SQ   SEQUENCE   194 AA;  20361 MW;  EDC8E0F602E02DC9 CRC64;


  Query Match            100.0%;  Score 52;  DB 2;  Length 194;
  Best Local Similarity  100.0%;  Pred. No. 0.077;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 PVVAESPKKP 10
          ||||||||||
Db    185 PVVAESPKKP 194


RESULT  12
Q9S3P4
ID   Q9S3P4      PRELIMINARY;      PRT;   194 AA.
AC   Q9S3P4;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C (FRAGMENT).
GN   OSPC.
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JSB;
RX   MEDLINE=96296448; PubMed=8709845;
RA   Livey I., Gibbs C.P., Schuster R., Dorner F.;
RT   "Evidence for lateral transfer and recombination in OspC variation in
RT   Lyme disease Borrelia.";
RL   Mol. Microbiol. 18:257-269(1995).
DR   EMBL; L42883; AAB36991.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
FT   NON_TER       1      1
FT   NON_TER     194    194
SQ   SEQUENCE   194 AA;  20446 MW;  CEEDC9FA5DF0D68F CRC64;


  Query Match            100.0%;  Score 52;  DB 2;  Length 194;
  Best Local Similarity  100.0%;  Pred. No. 0.077;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 PVVAESPKKP 10
          ||||||||||
Db    185 PVVAESPKKP 194


RESULT  13
P94229
ID   P94229      PRELIMINARY;      PRT;   194 AA.
AC   P94229;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OUTER SURFACE PROTEIN C (FRAGMENT).
GN   OSPC.
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=H9;
RX   MEDLINE=96296448; PubMed=8709845;
RA   Livey I., Gibbs C.P., Schuster R., Dorner F.;
RT   "Evidence for lateral transfer and recombination in OspC variation in
RT   Lyme disease Borrelia.";
RL   Mol. Microbiol. 18:257-269(1995).
DR   EMBL; L42888; AAB36996.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
FT   NON_TER       1      1
```

```
FT   NON_TER      194    194
SQ   SEQUENCE   194 AA;  20270 MW;  A016BB5336A9C981 CRC64;


  Query Match             100.0%;  Score 52;  DB 2;  Length 194;
  Best Local Similarity   100.0%;  Pred. No. 0.077;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 PVVAESPKKP 10
           ||||||||||
Db     185 PVVAESPKKP 194


RESULT  14
Q45030
ID   Q45030      PRELIMINARY;      PRT;   194 AA.
AC   Q45030;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   OSPA PROTEIN (FRAGMENT).
GN   OSPA.
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PLE;
RX   MEDLINE=95395018; PubMed=7665660;
RA   Jauris-Heipke S., Liegl G., Preac-Mursic V., Roessler D., Schwab E.,
RA   Soutschek E., Will G., Wilske B.;
RT   "Molecular analysis of genes encoding outer surface protein C (OspC)
RT   of Borrelia burgdorferi sensu lato: relationship to ospA genotype and
RT   evidence of lateral gene exchange of ospC.";
RL   J. Clin. Microbiol. 33:1860-1866(1995).
DR   EMBL; X80255; CAA56548.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
FT   NON_TER       1      1
FT   NON_TER     194    194
SQ   SEQUENCE   194 AA;  20523 MW;  11D409DBBFD23288 CRC64;


  Query Match             100.0%;  Score 52;  DB 2;  Length 194;
  Best Local Similarity   100.0%;  Pred. No. 0.077;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 PVVAESPKKP 10
           ||||||||||
Db     185 PVVAESPKKP 194


RESULT  15
Q45187
ID   Q45187      PRELIMINARY;      PRT;   207 AA.
AC   Q45187;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE   OUTER SURFACE PROTEIN C PRECURSOR.
GN   OSPC.
OS   Borrelia garinii.
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=29519;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PTROB;
RX   MEDLINE=95213332; PubMed=7699024;
RA   Wilske B., Jauris-Heipke S., Lobentanzer R., Pradel I.,
RA   Preac-Mursic V., Rossler D., Soutschek E., Johnson R.C.;
RT   "Phenotypic analysis of outer surface protein C (OspC) of Borrelia
RT   burgdorferi sensu lato by monoclonal antibodies: relationship to
RT   genospecies and OspA serotype.";
RL   J. Clin. Microbiol. 33:103-109(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PTROB;
RX   MEDLINE=95395018; PubMed=7665660;
RA   Jauris-Heipke S., Liegl G., Preac-Mursic V., Roessler D., Schwab E.,
RA   Soutschek E., Will G., Wilske B.;
RT   "Molecular analysis of genes encoding outer surface protein C (OspC)
RT   of Borrelia burgdorferi sensu lato: relationship to ospA genotype and
RT   evidence of lateral gene exchange of ospC.";
RL   J. Clin. Microbiol. 33:1860-1866(1995).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PSCF, PBAEII, PFIN, PFLK, PMUE, AND PSH;
RA   Marconi R.T., Hohenberger S., Jauris-Heipke S., Schulte-Spechtel U.,
RA   Lavoie C.P., Roessler D., Wilske B.;
RT   "Genetic analysis of B.garinii OspA-serotype 4 strains associated with
RT   neuroborreliosis: evidence for extensive genetic homogeneity.";
RL   Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; X83554; CAA58544.1; -.
DR   EMBL; AJ236908; CAB46238.1; -.
DR   EMBL; AJ132793; CAB46231.1; -.
DR   EMBL; AJ132796; CAB46234.1; -.
DR   EMBL; AJ132797; CAB46235.1; -.
DR   EMBL; AJ132798; CAB46236.1; -.
DR   EMBL; AJ236907; CAB46237.1; -.
DR   InterPro; IPR001800; Lipoprotein_6.
DR   Pfam; PF01441; Lipoprotein_6; 1.
DR   ProDom; PD001149; Lipoprotein_6; 1.
KW   Signal.
FT   SIGNAL        1     18       POTENTIAL.
FT   CHAIN        19    207       POTENTIAL.
SQ   SEQUENCE   207 AA;  22321 MW;  20889AA76E63A49D CRC64;


  Query Match             100.0%;  Score 52;  DB 2;  Length 207;
  Best Local Similarity   100.0%;  Pred. No. 0.082;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 PVVAESPKKP 10
           ||||||||||
Db     198 PVVAESPKKP 207


Search completed: October 12, 2002, 20:47:11
Job time: 11965 sec
```